Table of Contents for Draft SAB C-VPESS Report "Toward an Integrated and Expanded Approach for Ecological Valuation"

PART 1: OVERVIEW 6

1. INTRODUCTION 6
2. AN OVERVIEW OF KEY CONCEPTS 10
   - *2.1. The Concept of Ecosystem Services 10*
   - *2.2. The Concept of Value 12*
   - *2.3. The Concept of Ecological Valuation 16*
   - *2.4. Some Caveats Regarding Valuation 18*
3. ECOLOGICAL VALUATION AT EPA 21
   - *3.1. Policy Contexts at EPA Where Ecological Valuation Can be Important 21*
   - *3.2. Institutional and Other Issues Affecting Valuation at EPA 24*
   - *3.3. An Illustrative Example of Ecosystem Benefit Assessment at EPA 27*
4. AN INTEGRATED AND EXPANDED APPROACH TO ECOSYSTEM VALUATION: KEY FEATURES 31
5. IMPLEMENTING THE INTEGRATED AND EXPANDED APPROACH 34
   - *5.1. Early Consideration of Effects that are Socially Important 36*
   - *5.2. Challenges in Predicting Ecological Changes 37*
   - *5.3. Approaches for Characterizing Values 40*
   - *5.4. Communicating Results 43*
6. SUMMARY 44

PART 2: APPLYING THE APPROACH 46

1. INTRODUCTION 46
2. IMPLEMENTING THE CONCEPT OF ECOSYSTEM SERVICES 47
3. PREDICTION OF ECOLOGICAL EFFECTS 53
   - *3.1. Identifying the Stressors 55*
   - *3.2. Identifying Relevant Assessment Endpoints 55*
   - *3.3. The Use of Ecological Models 58*
   - *3.4. Gap between ecological models and the needs of ecological valuation 61*
   - *3.5. Closing the Gap 62*
4. INTRODUCTION TO DIFFERENT TYPES OF METHODS 71
   - *4.1. Biophysical Ranking Methods 71*
   - *4.2. Social-Psychological Methods 72*
   - *4.3. Economic Methods 74*
   - *4.4. Deliberative Processes that Involve Values 76*
   - *4.5. Public and Group Expressions of Value 78*
   - *4.6. Methods Using Cost as a Proxy for Value 79*
5. ECONOMIC VALUATION FOR NATIONAL RULEMAKING 81
   - *5.1. Introduction 81*
   - *5.2. Implementing the Proposed Approach 83*
   - *5.3. Conclusions and recommendations 99*
6. VALUATION FOR SITE-SPECIFIC DECISIONS 109
   - *6.1. Introduction 109*
   - *6.2. Opportunities for using valuation to inform contaminated property decisions 110*
   - *6.3. Use of source examples to illustrate recommendations 115*
   - *6.4. Recommendations 116*
   - *6.5. Source Example Analysis 118*
7. VALUATION IN COLLABORATIVE PARTNERSHIPS 146
8. ANALYSIS AND REPRESENTATION OF UNCERTAINTIES IN ECOLOGICAL VALUATION AND COMMUNICATION OF ECOLOGICAL VALUATION INFORMATION 147

8.1. Uncertainty *147*
8.2. Communication and valuation *164*
9. RECOMMENDATIONS 173
9.1. Research *173*
9.2. Guidance documents *173*
9.3. Institutional Recommendations *173*

PART 3: METHODS FOR IMPLEMENTING APPROACH 174

1. INTRODUCTION 174
2. BIO-PHYSICAL RANKING METHODS 175
2.1. Introduction *175*
2.2. Conservation Value Method *175*
2.3. Rankings Based on Energy and Material Flows *188*
3. SOCIO-PSYCHOLOGICAL APPROACHES 204
3.1. Introduction *204*
3.2. Brief Description of the Methods *206*
3.3. Relation of Methods to the CVPESS Expanded and Integrated Assessment Framework *213*
3.4. Transfer of Socio-Psychological Information *226*
4. ECONOMIC METHODS 227
4.1. Overview *227*
4.2. Market-Based Methods *230*
4.3. Non-market Methods – Revealed Preference *232*
4.4. Non-market Methods – Stated Preference *240*
4.5. Ecosystem Benefit Indicators *247*
4.6. Transfer of Economic Information *250*
5. DELIBERATIVE PROCESSES FOR ELICITING VALUES 251
5.1. Valuation by Decision Aiding/Structured Decision Making *251*
5.2. Mediated Modeling *261*
6. PUBLIC AND GROUP EXPRESSIONS OF SOCIAL/CIVIC VALUE 270
6.1. Introduction *270*
6.2. Referenda and Initiatives *271*
6.3. Citizen Valuation Juries *282*
7. METHODS USING COST AS A PROXY FOR VALUE 291
7.1. Replacement Costs *291*
7.2. Tradeable permits *292*
7.3. Habitat Equivalency Analysis *293*

APPENDIX A: SURVEY ISSUES FOR ECOLOGICAL VALUATION: CURRENT BEST PRACTICES AND RECOMMENDATIONS FOR RESEARCH 298

Defining Survey Research *298*
Contemporary Uses of Surveys *298*
Evidence of Survey Accuracy *299*
Reasons to Conduct Surveys *301*
Designs of Surveys *302*
Total Survey Error *306*
Sampling *307*
Nonresponse Error *310*
Question Design *312*
Mode of Data Collection *318*
General Conclusions About the Use of Survey Methods *322*
Challenges in Using Surveys For Ecosystem Protection Valuation *323*
Addressing the question: "People do not hold stable, well-defined tradeoffs for the component attributes of ecosystem valuation. Methods, surveys or other, will inevitably shape the construction of those values." What are the practical and legal implications of this? *339*

More generally, when and why should EPA survey the public about their values, desires, preferences? Addressing the questions: "The public likely hold values that arguably should not be part of a normative, regulatory decision (what should be done) -- where, in EPA policy processes, might surveys be used? Advantages/disadvantages at each level *351*

APPENDIX B: REFERENCES **359**

APPENDIX C: ENDNOTES **383**

Lists of Tables, Text Boxes, and Figures

Tables

Table 1: Usage of Terms 15
Table 2: Table of Generic Assessment Endpoints Reproduced from U.S.EPA, 2003 56
Table 3: Farber et al., 2006), Classification of Benefits Transfer of Ecosystem Services from One Context to Another 65
Table 4: Impacts of Climate Change on Stream Ecosystems (Source: Kling, et al., 2003, "Confronting Climate Change in the Great Lakes Region: Impacts on our Communities and Ecosystems," Union of Concerned Scientists, Cambridge, MA, and Ecological Society of America, Washington, D.C., p. 26 90
Table 5: Steps in the Superfund Process 111
Table 6: Ecosystem Service Matrix for Leviathon Mine (from Wilson, 2004) 127
Table 7: Steps to Execute the Conservation Value Method and Application of the Steps Outside the EPA 181
Table 8: Examples of Federal Surveys 300

Text Boxes

Text Box 1: The Aquaculture Effluent Guidelines 100
Text Box 2: The CAFO Effluent Guidelines 102
Text Box 3: The Prospective Benefits of the Clean Air Act Amendments 104
Text Box 4: Charles George Landfill 141
Text Box 5: Dupage County Landfill 142
Text Box 6: Avtex Fibers Site 142
Text Box 7: Leviathan Mine Superfund Site 143
Text Box 8: Table of Examples of Measures from Different Examples Discussed in this Report 168
Text Box 9: What Are Conjoint Surveys of Attitudes? 209
Text Box 10: The Constructive Nature of Environmental Preferences 251
Text Box 11: Structuring the Valuation Problem 253
Text Box 12: How to Monetize if You Must 255
Text Box 13: Direct Analysis of Public Decisions to Accept Pollution or Resource Depletion 271
Text Box 14: Referendum/Initiative Analysis Followed by a Survey 272
Text Box 15: Public Decisions to Accept Pollution or Resource Depletion Followed by a Survey 273
Text Box 16: Referenda and Initiatives Used to Validate Contingent Valuation 279
Text Box 17: AValuation Exercise Illustrating Use of Citizen Juries 284
Text Box 18: Other Valuation Exercises Using Citizen Juries 285
Text Box 19: Other Group-Based Approaches to Valuation 285

Figures

Figure 1: Components of Ecological Valuation 16
Figure 2: Process for Implementing an Expanded and Integrated Approach to Ecological Valuation 35
Figure 3: Illustration of How Abundance of Functional Groups Can Characterize the Complexity of an Ecosystem 50
Figure 4: Illustration from Covich et al., 2004, Showing Relationships of Major Functional Types to Ecological Services 51
Figure 5: General Overview of the Impact of CAFOs 84
Figure 6: Changing Focus from Remediation to Redevelopment Would Benefit from Increased Integration of Valuation Analysis with Traditional Process Steps 112
Figure 7: Ecological Services Linkages Conceptual Model 122

PART 1: OVERVIEW

1. INTRODUCTION

EPA's Mission Regarding Ecosystem Protection. The Environmental Protection Agency's (EPA's) mission is to protect human health and the environment. During its history, the EPA has focused decision making and much of its expertise on the first part of this mission, in particular the risks to human health from chemical stressors in the environment. Although protecting human health is the bedrock of the EPA's traditional expertise, the broad mission of the EPA goes beyond this. In fact, EPA's Strategic Plan explicitly identifies the need to ensure "healthy communities and ecosystems" as one of its five major goals (U.S Environmental Protection Agency 2003) and EPA's efforts in protecting ecological resources--and its authority for doing so--have been documented in Agency publications and independent historical sources (U.S. Environmental Protection Agency 1994); U.S. Environmental Protection Agency Risk Assessment Forum 2003, (U.S. Environmental Protection Agency Science Advisory Board 2000), (Hays 1989); (Russell III 1993).

EPA's mission to protect the environment requires attention to ecological systems to ensure the wise and thoughtful use and protection of our environment. An "ecosystem" is the term used by ecologists to describe living organisms plus their physical environment working together. For example, a forest ecosystem is comprised of the trees in the forest plus the birds, insects, soil micro-organisms, and streams that inhabit or run through it. Ecosystems provide basic life support for human and animal populations and are the source of spiritual, aesthetic and other human experiences that are valued in many ways by many people.

Given the important role that ecosystems play in our lives, changes in the state of these systems or the flow of services they provide can have important implications. Many EPA actions (e.g., regulations, rules, programs, policy decisions) affect the condition of the environment and the flow of ecological services from it. EPA actions can lead to improvement or deterioration of ecosystems or prevent degradation that

would otherwise have occurred. These impacts can occur both at a relatively small, local scale as well as more broadly at a national scale.

Despite their importance, to date, ecological impacts have received relatively limited consideration in EPA policy analyses. EPA's ecological analysis has generally focused primarily on ecological endpoints such as those identified by tests required for pesticide regulation (e.g., effects on survival, growth, and reproduction of aquatic invertebrates, fish, birds, mammals, and both terrestrial and aquatic plants) or mortality to fish, birds, and plants and, more generally, animals, wildlife, aquatic life, as required by provisions of several laws[1] administered by the Agency (U.S. Environmental Protection Agency Risk Assessment Forum 2003). However, given EPA's responsibility to ensure healthy communities and ecosystems, the Agency's actions must encompass the key structural and functional characteristics of communities and ecosystems, not simply endpoints related to population impacts. Failure to consider ecological impacts as fully as possible can lead to distorted policy decisions. This can occur, for example, when actions are evaluated based primarily on their impacts on human health, without adequate recognition of potentially important ecosystem impacts.

In addition to its mission to protect ecosystems, EPA also seeks to evaluate policy options and make policy decisions with a recognition of the tradeoffs that are inevitably involved. To promote good decision-making, policy makers require information about how ecosystems contribute to society's well-being and how those contributions are valued. This need is increasingly recognized both within and outside the Agency. The stated goal of EPA's recently released *Ecological Benefits Assessment Strategic Plan (EBASP)* is to "help improve Agency decisionmaking by enhancing EPA's ability to identify, quantify, and value the ecological benefits of existing and proposed policies" (p. xv). In addition, information about the value of ecosystems and the associated impacts of EPA actions can help inform the public about the need for ecosystem protection and the extent to which specific policy alternatives address that need.

Despite EPA's stated mission and mandates, there is a gap between the need for protection of ecological systems and services and EPA's ability to address this need.

This report is a step toward filling that gap. It describes how an integrated and expanded approach for valuation of ecological systems and services can help the Agency describe and measure the value of protecting ecological systems and services and hence better meet its overall mission.

This report was prepared by the Committee on Valuing the Protection of Ecological Systems and Services (C-VPESS), which was formed by EPA's Science Advisory Board (SAB). The SAB saw a need to complement the Agency's ongoing work in ecological science, ecological risk assessment, and ecological benefit assessment by offering advice on how EPA might better value the protection of ecological systems and services and how that information might better support decision making to protect ecological resources. Toward this end, it formed C-VPESS,[2] an interdisciplinary group of experts from the following areas: decision science, ecology, economics, engineering, philosophy, psychology, and social sciences with emphasis on ecosystem protection.[3] C-VPESS began its work in 2003 on a project designed to strengthen the Agency's analysis for protecting ecological resources. The purpose of science advice on ecological valuation is to strengthen the Agency's knowledge and set of analytical tools to help navigate difficult trade-offs that inevitably arise when regulatory or other decisions must be made to protect ecological resources. In this project the SAB set the goals of: a) assessing Agency needs and the state of the art and science of valuing protection of ecological systems and services and b) identifying key areas for improving knowledge, methodologies, practice, and research at EPA.

<u>Scope of report and intended audience</u>. This report provides advice for strengthening the Agency's approaches for valuing the protection of ecological systems and services, facilitating their use by decision makers, and identifying the key research areas needed to strengthen the science base.[4] It focuses on the need for an expanded and integrated approach for valuing EPA's efforts to protect ecological systems and services. It provides advice to the Administrator, EPA managers, EPA scientists and analysts, and EPA staff across the Agency concerned with ecological protection. It adopts a broad view of EPA's work, which it understands to encompass national rulemaking, regional decision making, and programs in general that protect ecological systems and services. It

outlines a call for EPA to expand and integrate its approach in important ways.

This report appears at a time when there is lively interest internationally, nationally, and at EPA itself in the issue of valuing the protection of ecological systems and services. Since the establishment of the SAB C-VPESS major reports have been developed by others focusing on how to improve the characterization of the important role of ecological resources (Millennium Ecosystem Assessment Board 2003; Silva and Pagiola 2003; National Research Council 2004; Pagiola, von Ritter et al. 2004; Millennium Ecosystem Assessment 2005). In addition, the Agency itself has engaged in efforts to improve ecological valuation. The most recent product of these efforts is the *EBASP* report noted above (U.S. Environmental Protection Agency 2006). This report discusses in length past and current EPA efforts to improve ecological valuation (see Appendix B), which have focused on economic valuation for use in benefit-cost analysis. EPA has also sought to strengthen the science supporting ecological valuation through the extramural Science to Achieve Results (STAR) grants program. STAR grants involving ecological valuation have primarily applied economic valuation methods to various ecosystem services.

The committee's work has benefited from and has built upon these recent efforts. The C-VPESS distinguishes its work from those efforts, however, in the following ways. First, the C-VPESS focuses on EPA as an audience for its work. In particular, it focuses on how EPA can value its own contributions to the protection of ecological systems and services, so that the agency can make better decisions in its eco-protection programs. Many of the recent studies (for example, the Millennium Assessment and NRC report) do not consider the specific policy contexts or constraints faced by EPA. Second, most previous work has focused on economic valuation. In contrast, C-VPESS is inter-disciplinary and does not focus solely on economic methods or values. The committee will offer advice on several approaches to characterizing or estimating values and in each case will emphasize issues relevant to EPA policy and decision-making and address how the Agency could better represent the value of ecological protection.

2. AN OVERVIEW OF KEY CONCEPTS

2.1. The Concept of Ecosystem Services

The term "ecosystem" describes a dynamic complex of plant, animal, and microorganism communities and the non-living environment, interacting as a unit. For example, ecosystems can describe organism-physical environment interactions in a woodlot, a watershed, or a larger landscape. Ecosystems encompass all organisms within the prescribed area, including humans, who are often the dominant element. Ecosystem "functions" or "processes" are the characteristic physical, chemical, and biological activities that influence the flows, storage, and transformation of materials and energy within and through ecosystems (U.S Environmental Protection Agency 2004). These include processes that link organisms with their physical environment (e.g., primary productivity and the cycling of nutrients and water) and processes that link organisms with each other, indirectly influencing flows of energy, water and nutrients (e.g., such as pollination, predation and parasitism). These processes in total describe the functioning of ecosystems.

"Ecosystem services" is an anthropocentric concept denoting the contributions that ecosystems make to people either directly or indirectly. In popular terminology these contributions are sometimes labeled the "benefits" that humans derive from ecosystems.[5] The following operational categorization of ecosystem services has recently been proposed by the Millennium Ecosystem Assessment:

a) Provisioning services (products obtained from ecosystems). These include food, fuelwood, fiber, biochemicals, genetic resources and fresh water. Generally these services are traded in the open marketplace.

b) Regulating services (benefits received from regulation of ecosystem processes). This category includes a host of pathways that stem from the presence and functioning of ecosystems and influence people in positive ways, both direct and indirect. These include flood protection, human disease regulation, water purification, air quality maintenance, pollination,

pest control and climate control. These services are generally not marketed but many have clear value to society and this value will increase for many of these services as the many dimensions of global change proceed.

c) Cultural services (the nonmaterial benefits people obtain from ecosystems). Ecosystems contribute to the cultural, spiritual and aesthetic dimensions of people's well-being. They also contribute to establishing a sense of place.

d) Supporting services. These are the processes that maintain ecosystem functioning such as: soil formation, primary productivity, biogeochemistry, and provisioning of habitat. They all affect human well-being, but generally indirectly through their support of the provisioning, regulating and cultural service functions.

This categorization suggests a very broad definition of services, limited only by the requirement of a contribution (direct or indirect) to human well-being. This broad approach reflects the need to recognize the myriad ways in which ecosystems support human life and contribute to human well-being. [6] Alternatively, ecosystem services can be defined more narrowly to include only end-point services that contribute directly to well-being. (See Part 2 of this report for a more detailed discussion of the implications of using a broad vs. narrower definition.)

Whether broadly or narrowly defined, the concept of ecosystem services provides an approach to evaluating the many ways in which ecological changes induced by human actions affect human well-being. However, ecosystems can be valued for reasons that are independent of human well-being. As discussed below, the committee recognizes that ecosystems can be important not only because of the services they provide directly or indirectly but also for other non-anthropocentric reasons, including respect for nature based on moral, religious, or spiritual beliefs and commitments.

2.2. The Concept of Value

Because only people define values, all values are anthropogenic. However, there is disagreement about whether the ultimate goal of all values is only to promote human well-being, so there is disagreement about whether all values are anthropocentric. When people talk about environmental values, the values of nature, or the values of ecological systems and services, they may have different things in mind. People have moral, economic, religious, aesthetic, and other interests, all of which can affect their thoughts, attitudes, and actions toward nature in general and, more specifically, toward ecosystems and the services they provide.

The most basic philosophical distinction in values is the distinction between means and ends. To value something as a means is to value it for its usefulness in helping to realize or bring about some thing or state of affairs that is valued in its own right or as an end. Things valued for their usefulness as means in this sense are said to have instrumental value. Alternatively, something can be valued for its own sake or as an end. Things valued as ends are sometimes said to have intrinsic value. [7] If intrinsic value applies to things other than human beings or human experiences, then this conception of value is non-anthropocentric. Some people defend a non-anthropocentric conception of value or goodness (Goodpaster 1978; Taylor 1986; Rolston III 1991). However, others argue that only human beings or human experiences have intrinsic value, thereby defending an anthropocentric conception of intrinsic value (Sidgwick 1901; Glover 1984; Williams 1994).

Some people also claim that the very "existence" of a species or ecological system has value in addition to any instrumental value derived from the usefulness of the services it provides. This claim can mean several different things. If it means that the existence of an ecological system is valuable because people derive satisfaction from its existence independent of specific uses they may make of its services, then it has what economists call "existence value." This concept is anthropocentric. In addition, this value can been viewed as a kind of instrumental value, since it is based on the premise that the existence of the species or ecological system is one of many things that generate human satisfaction, and that the various things that contribute to human satisfaction are

potentially substitutable or commensurable. In contrast, some people claim that an ecological system may have intrinsic value of its own, and hence its existence is valuable for its own sake. If the explanation of this claim refers to reasons that are independent of the contribution that the existence of the ecological system can make to human well-being, then as noted above this is a claim that the ecosystem or one or more of its components has a non-anthropocentric intrinsic value.

This committee recognizes that there are many possible sources of value derived from ecosystems and the services they provide. To reflect this, throughout this report, the term "value" is used broadly to include values that stem from contributions to human well-being as well as values that reflect other considerations, such as social and civil norms (including rights) and moral, religious, and spiritual beliefs and commitments. (See Table 1 for a summary of the usage of key terms throughout this report.)

As distinct from the broader concept of value, in this report we use the term “benefit” to refer more narrowly to the contribution of ecosystems and their services to human well-being. As such, benefits include only anthropocentric sources of value. However, as defined here, the term “benefits” includes both the economic/utilitarian concept of benefits based on individuals’ preferences and benefits based on communitarian or constructed preferences. In addition, throughout the report benefits are defined relative to changes in the state of an ecosystem or the flow of services it provides stemming from an actual or proposed action by EPA. Thus, the term “ecosystem benefits” refers to the positive contribution to human well-being of a change in an ecological system and/or its services. A negative contribution, for example from damages to an ecosystem, can be viewed as a “negative benefit” or cost. Similarly, the term “valuation” will refer to the process of estimating or measuring either the value of, or the value of a change in, an ecosystem, its components, or the services it provides.

Finally, throughout the report the concept of value or benefit refers to social value or benefit of a given change, which could differ from the value or benefit to a private party such as a firm. For example, allowing a firm to emit one more unit of a pollutant can have a positive value or benefit for that firm (equal to the savings in abatement costs)

but this private value is not the social value of that increase (which might very well be negative, i.e., entail a net social cost).

Table 1: Usage of Terms

For purposes of this report, the following terms are used as indicated:

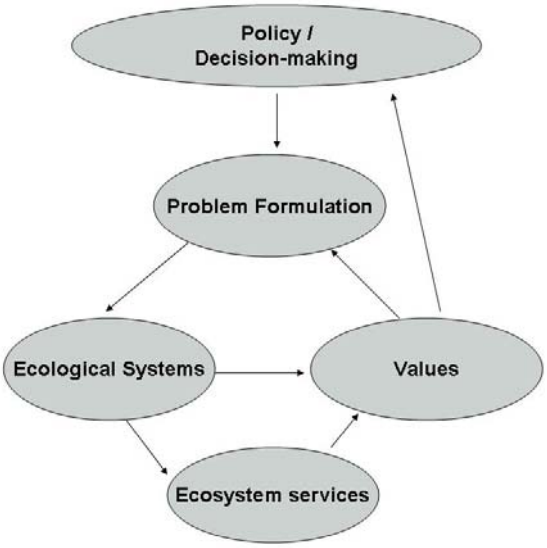
Ecosystem: A dynamic complex of plant, animal, and microorganism communities and the non-living environment, interacting as a unit.
Ecosystem functions or processes: The characteristic physical, chemical, and biological activities that influence the flows, storage, and transformation of materials and energy within and through ecosystems. These include processes that link organisms with their physical environment, (e.g., primary productivity and the cycling of nutrients and water) and processes that link organisms with each other (e.g., pollination, predation, and parasitism).
Ecosystem Services: Those ecological characteristics, functions or processes that directly or indirectly contribute to the well-being of human populations or have the potential to do so in the future.
Value: This term is used broadly to include well-being; social and civil norms, including rights; and moral, religious, aesthetic, and spiritual beliefs and commitments.
Valuation: The process of estimating or measuring either the value of, or the value of a change in, an ecosystem, its components, or the services it provides.
Valuation Method: A methodology, based on theory and data, for estimating or measuring the value of, or the value of a change in, an ecosystem, its components or the services it provides.
Monetary Valuation: Valuation in which estimates are expressed in monetary units.
Non-monetary Valuation: Valuation in which estimates are expressed in non-monetary terms.
Benefits: The contribution of ecosystems and their services to human well-being and satisfaction. This includes both the economic/utilitarian concept of benefits based on individuals' preferences and benefits based on communitarian or constructed preferences.
Economic Valuation Methods: Methods that estimate the tradeoffs individuals are willing to make for improvements in, or to avoid degradation of, an ecosystem, its components, or the services it provides. These approaches typically focus on the amount of money an individual is willing to forgo or pay to enjoy a particular positive change (willingness-to-pay) or the amount of monetary compensation a person would accept in lieu of receiving that change (willingness to accept). This includes benefits derived from both use and non-use values, and depends on both the individual's income and current prices.
Social-Psychological Valuation Methods: Methods that focus on individuals' judgments of the relative importance of, acceptance of or preferences for changes in ecosystems, their components, or the services they provide, typically focusing on choices or ratings among alternatives. Individuals making the judgments may respond on their own behalf or on behalf of others (society at large or specified sub-groups) and the basis for judgments may be changes in individual welfare, or civic or ethical/moral obligations relevant to ecosystems and ecosystem services.

.

2.3. The Concept of Ecological Valuation

"Valuation" is the process of estimating or measuring either the value of, or the value of a change in, an ecosystem, its components, or the services it provides. The committee is focusing on valuation in EPA contexts where there is an environmental protection decision to be made. The major components of ecological valuation are depicted in Figure 1.

Figure 1: Components of Ecological Valuation



Some of the components of ecological valuation parallel components of the ecological risk framework that underlies the ecological risk guidelines developed by EPA to support decision making to protect ecological resources (U.S. Environmental Protection Agency Risk Assessment Forum 1992; U.S. Environmental Protection Agency Risk Assessment Forum 1998). The committee views ecological valuation as a complement to ecological risk assessment. Both begin with an EPA decision or policy context for which information about ecological effects is needed. This leads to a formulation of the problem and identification of the purpose and objectives of the analysis and the policy options that will be considered. In addition, both ecological risk

assessment and ecological valuation involve prediction and estimation of possible ecological effects of the EPA action or decision that is under consideration.

However, ecological valuation goes beyond ecological risk assessment in an important way. Risk assessments typically focus on predicting the magnitudes and likelihoods of possible adverse effects on species, populations, locations, etc., but do not provide information about the societal importance or significance of these effects. In contrast, as depicted in Figure 1, ecological valuation takes the predicted ecological effects and seeks to characterize their importance to society by providing information on the value society places on the ecological improvements or the loss they experience from ecological degradation. As noted above, these values can reflect either changes in the flow of services provided by the ecosystem or values that are attached directly to the ecosystem itself that are independent of its contribution to human well-being. Information about these values is an important input in the evaluation of alternative decisions or policy options.

There are a number of methods that can be used for estimating or measuring values. In some cases, these methods lead to estimates that are expressed in monetary units ("monetary valuation"), while in other cases estimates are expressed in non-monetary terms ("non-monetary valuation"). In addition, these methods differ in their focus and, in some cases, their underlying premises. For example, economic valuation measures benefits by estimating the tradeoffs individuals are willing to make for ecological improvements or to avoid ecological degradation. These approaches typically (but not always) estimate benefits, including existence values, in monetary units that measure the amount of money an individual is willing to pay to enjoy a particular positive change (willingness-to-pay) or the amount of monetary compensation a person would accept in lieu of receiving that change (willingness to accept) given the individual's current income and the current prices of other goods and services that can be purchased with that income.[8]. Alternatively, social-psychological methods for valuation focus on individuals' judgments of the relative importance of, acceptance of, or preferences for ecological changes. These approaches typically focus on choices or ratings among sets of alternative policies, and may include comparisons with potentially

competing social and economic goals. Individuals making the judgments may respond on their own behalf or on behalf of others (society at large or specified sub-groups) and the basis for judgments may be changes in individual well-being, or civic or ethical/moral obligations relevant to ecosystems and ecosystem services. Similarly, assessment methods based on voting or other group expressions of social/civic values provide information about human values revealed through these processes

2.4. Some Caveats Regarding Valuation

A basic tenet of valuation as defined in this report is that it seeks information about the value of protecting ecological systems and services as expressed by lay individuals within society. The presumption is that, in a democratic society, the values held by individuals within that society should be considered in public policy decisions. The involvement of citizens in decisions about their future environments and what would best serve their individual and collective well-being is presumably a basic tenant of democratic societies. In addition, while under this premise everyone's values count, values held, expressed and advocated by larger numbers of people should carry more weight in public policy.

However, some (committee members??) believe that, for complex problems such as ecosystem protection, majority values or values held by the public are not an appropriate basis for public policy decisions (for reasons discussed below). In such cases, they suggest that it would be better to leave the public out of the process altogether and have experts (e.g., ecologists, biologists, toxicologists) determine ecosystem and ecosystem service protection policies. However, it is not always clear what the necessary normative principles are for such decisions, or who should decide among competing principles. In some cases, there may be high levels of agreement among experts about the bio-physical outcomes of proposed policies, and even about the implications of those outcomes for individual and social well-being. When expert judgments on these matters are incongruent with the beliefs, preferences or intentions of the public, there is a need to determine why, and what information, deliberations or other interventions might be needed to achieve better congruence of expert and public opinion.

Concerns about basing policy decisions on values expressed by individuals rather than expert judgments stem from at least two sources. The first is a view that the preferences that people express are not well-formed or stable and are easily subject to (intentional or unintentional) manipulation (see a detailed discussion in Appendix ??). This suggests that expressed attitudes and preferences can be changed if the judgment an individual is asked to make is presented in a different way. For example, studies have shown that responses to surveys sometimes exhibit instability in the form of response choice order effects, question order effects, question framing effects and interviewer effects. However, most of these effects, although statistically significant in some studies, have been quite small. In addition, many studies that seem to suggest that people's opinions are uncrystallized and easily manipulated have been shown to have had problems with the research design, such as problematic choice of the participant population involved or the setting in which the data were collected. Nonetheless, for contexts that are very unfamiliar, initial responses to survey questions regarding attitudes and values may not reflect well-formed preferences. Thus, if not carefully conducted, valuation studies based on techniques such as surveys could yield misleading representations of the public's preferences.

The second source of concern about the use of values from the public (rather than experts) in policy choice relates to the quantity and quality of information individuals have when expressing their values. In principle, values used in public policy decisions should reflect all available information. However, for complex issues such as ecosystem protection, individuals will generally not have the same level of information about, or understanding of, the related science as experts, and the values they express either through survey responses or through their behavior will reflect that incomplete information. Some argue that if members of the public are in fact not well informed, perhaps they cannot offer thoughtful opinions or make rational choices about related policy matters, so allowing public influence on decisions bearing on those issues would be irresponsible (de Tocqueville, 1835; Schumpeter, 1950). At a minimum, it is clear that expressed values will depend on the information that individuals have available, and more complete information will lead to more accurate value assessments.

The discussion above suggests that policy makers using values expressed by the public as an input into decisionmaking should be cognizant of the above concerns and interpret the value estimates accordingly. *Ceteris paribus*, policy-makers should put more weight on measures of public preferences that are based on well-informed and thoughtful expressions of value. Additionally, EPA should consider taking direct steps to assess the level of understanding brought to issues that have complex policy and scientific implications and the implications for valuation, particularly where there are concerns about the amount of information the public understands about the issues addressed by the Agency. [9]

3. ECOLOGICAL VALUATION AT EPA

There are several different contexts in which EPA policy decisions result in ecological impacts and hence in which the need for ecological valuation will arise. In addition, EPA operates within a set of institutional, legal, organizational and practical constraints that affect this process at the Agency. Thus, EPA has specific needs in this regard that must be recognized and addressed. These needs arise in different parts of the Agency for different purposes and for different audiences. Some of the needs present structured requirements for valuing protection of ecological systems and services, while needs in other contexts are less prescriptive.

3.1. Policy Contexts at EPA Where Ecological Valuation Can be Important

There are at least three policy contexts in which information about the value of ecological systems and services could be very useful to EPA: a) national rule-making; b) regional decision-making; and c) local assessment and evaluation.

Benefit assessments are required for national rulemaking by two of EPA's governing statutes (the Toxic Substances Control Act and the Federal Insecticide, Fungicide and Rodenticide Act) and by Executive Order 12866 for "significant regulatory actions." The circular on "Regulatory Analysis" issued by the Office of Management and Budget (OMB) in September 2003, OMB Circular A-4, identified key elements of a regulatory analysis for such "economically significant rules." One of these elements is an evaluation of the benefits and costs of a proposed regulatory action and the main alternatives identified. The circular explicitly defines benefits using the economic/utilitarian concept of willingness to pay (or willingness to accept).[10] The circular contains general guidance on how to provide monetized, quantitative, and qualitative information to characterize benefits as fully as possible, and EPA itself has developed initial guidance for ecological benefit assessment (U.S. Environmental Protection Agency 2000, reference EBASP too?). In developing the *EBASP*, EPA identified the need for improved models and methods to help implement the requirements of the circular as they relate to ecological valuation. The Agency identified the need both to expand methods and data for economic valuation and to explore other assessment

methods to provide information on ecological effects that are currently not quantified or monetized and assigned an implicit value of $0. Managers seek approaches that are "sound, credible, and scientifically supportable" as well as flexible, affordable, and able to be implemented within the time constraints required by rulemaking (U.S. Environmental Protection Agency Science Advisory Board 2004 –check this quote…..).

EPA's regional offices, although generally not responsible for national rule-making, are responsible for several kinds of regional and local decisions and activities where the benefits of ecological protection are potentially important. These include:

- Priority setting for regional action, such as targeting projects for wetland restoration and enhancement or identifying critical ecosystems or ecological resources for regional attention;
- Setting Supplemental Environmental Protection (SEPs) penalties for enforcement cases where those penalties involve protection of ecological systems and services;
- Choice of options for Superfund and Resource Conservation and Recovery Act (RCRA) cleanups that could take ecological benefits into account;
- Review of Environmental Impact Statements prepared by other federal agencies to comply with the National Environmental Protection Act; and
- Assisting state and local governments and other federal Agencies with protecting lands and land uses, where assessment of the value of protection options could help decision-makers make better-informed decisions.

Regions also seek low-cost methods that can be implemented quickly to inform site-specific decisions. They seek methods that provide information on the value of ecological services; ecological diversity; conservation opportunities and threats; sustainability; and historical and cultural values associated with ecological systems or parts of ecosystems at the watershed or landscape scale. Regions experience the need to

communicate the value of ecological protection as they collaborate with other federal agencies and with government partners at the local, state, and regional levels.

The need to assess the ecological benefits of policy options arises in most of the Agency's decisions, including the assessment of ecological protection programs. EPA's need to assess the value of its ecological protection programs has two dimensions: 1) a retrospective dimension, because assessments focus on the value of EPA's current and past protection efforts, and 2) a prospective dimension, because such assessments are meant to inform decisions about future EPA programs and priorities. Program assessments are mandated for EPA, as they are for all agencies of the executive branch, by the Government Performance and Results Act (GPRA) of 1993. As part of that assessment, OMB requires EPA to periodically identify its strategic goals and describe both the social costs and budget costs associated with them. EPA's Strategic Plan for 2003-2008 described the current social costs and benefits of EPA's programs and policies under each strategic goal area for the year 2002 (U.S. Environmental Protection Agency 2003). This analysis repeatedly points out that EPA lacks data and methods to quantify the ecological benefits associated with the goals in its strategic plan.

In addition, GPRA established requirements for assessing the effectiveness of federal programs. Part of that assessment involves assessing the outcomes of programs intended to protect ecological resources. EPA must report annually on its progress in meeting program objectives linked to strategic plan goals and must engage periodically in an in-depth review [through the Program Assessment Rating Tool (PART)] of selected programs to identify their net benefits and to evaluate their effectiveness in delivering meaningful, ambitious program outcomes. Characterizing ecological benefits associated with EPA programs is a necessary part of the program assessment process.

Although ecological valuation can be an important part of program assessment, this report focuses on the use of valuation to inform policy decisions relating to national rule-making and regional and local priorities and activities. Nonetheless, the committee believes that the methods and issues discussed throughout the report can be used to improve the evaluation of EPA programs that protect ecosystems by demonstrating and, where possible, estimating the ecological benefits derived from those programs.

3.2. Institutional and Other Issues Affecting Valuation at EPA

The committee recognizes that ecological valuation at EPA must be conducted within a set of institutional, legal, organizational, and practical constraints that affect what is and can be done to incorporate ecosystem values into policy evaluations. These constraints include procedural requirements relating, for example, to timing and oversight, as well as the Agency's own resource constraints (both monetary and personnel). In an effort to better understand these issues and their implications for the committee's charge, the committee conducted a series of interviews with Agency staff. [11] The interviews were focused on the process of developing benefit analyses for Regulatory Impact Assessment (RIA) for rulemaking and the relationship between EPA and the Office of Management and Budget. However, many of the questions raised are equally applicable to strategic planning, performance reviews, regional analysis, and other situations in which the Agency is called upon to assess the value of ecosystems and the services they provide. Below are some key observations made by the committee based on those interviews.

EPA has a formal rule-development process with several stages, each of which imposes demands on the Agency, and the Agency also develops rules to meet court-imposed deadlines. However, despite the commonality of the underlying rule-development process, it is clear that there is no single way in which analysts within the Agency assess the tradeoffs that people would be willing to make to enhance ecosystems. Practices vary considerably across program offices, reflecting differences in mission, in-house expertise, etc. Program offices have different statutory and strategic missions. The organization, financing, and skills of the program offices differ enormously. The National Center for Environmental Economics (NCEE) is the Agency's centralized reviewer of economic analysis within the Agency. [12] However, the primary expertise and development of the rules resides within the program offices.

Secondly, the timing of the process largely determines the kinds of analytical techniques that are employed. This is related to court-imposed deadlines on the rule process, as well as intervening requirements related to the collection and analysis of new data. The scientific community is accustomed to much longer time horizons for their

analyses. Unfortunately, collecting new data poses a significant bureaucratic problem for the Agency. To collect original data, the Agency must submit an Information Collection Request (ICR), which is reviewed within the Agency and by OMB. This hurdle is required by the Paperwork Reduction Act and imposes the review responsibility on OMB. There is no discretion under the legislation in applying this review requirement (except in cases associated with the development of survey instruments). Nonetheless, the requirement can add a significant amount of time to the assessment process. With perhaps a year or two at most to conduct a study, this kind of review significantly limits the kind of analysis the Agency can conduct. In particular, it implies that the Agency must by necessity rely heavily on previous studies through the application of benefits transfer techniques.

A third issue is the role of the Office of Management and Budget (OMB) in defining or directing ecosystem valuation exercises at EPA. OMB acts as an oversight body that reviews EPA's benefit analyses. EPA is required to provide sufficient justification for its claims regarding the benefits of its actions, including any ecological benefits. As noted above, EPA has been given explicit guidance by OMB in the Circular A-4, which the committee views as a reasonable document on its own because of its call for a full characterization of the impacts of different policy options, including where possible a characterization of benefits that cannot be monetized or cannot be quantified (Office of Management and Budget 2003). For a benefit or cost that cannot be expressed in monetary terms, the Circular instructs Agency staff to "try to measure it in terms of its physical units," or, if this is not possible either, to "describe the benefit or cost qualitatively" (add page number). [13] Thus, although Circular A-4 does not require that all benefits be monetized, it does require at a minimum a scientific characterization of those benefits. However, little guidance is provided on how this should be done. Instead, the Circular urges regulators to "exercise professional judgment in identifying the importance of non-quantified factors and assess as best you can how they might change the ranking of alternatives based on estimated net benefits" (add page number).

In conducting benefit assessments, EPA has an incentive to use methods that have been accepted by OMB in the past. This creates a bias toward the status quo and a

disincentive to explore new or innovative approaches. The committee recognizes the value of consistency in the methods used for valuation, but also sees the limitations resulting from sole reliance on previously approved methods and the potential benefits from efforts to explore innovative or expanded approaches.

A related issue involves review of Regulatory Impact Assessments (RIAs) by external parties. The Agency does not take a standardized approach to RIA review. EPA staff and managers reported that peer review was focused only on "novel" elements of an analysis. This raises the question of how the Agency (and perhaps OMB) defines "novel." Moreover, the novelty standard actually creates a clear incentive to avoid conducting novel analyses (however defined). It is clearly cheaper and quicker to avoid review altogether. This suggests a possible role for a standing expert body that can bring consistency to the review of analysis, avoid duplication of review, and be sensitive to timing and resource constraints.

Finally, the committee notes the importance of the organization of assessment science within the Agency. Currently, the Agency relies upon a variety of offices to develop assessments, with varying degrees of reliance on other offices (e.g., NCEE) or outside assistance. It is not clear which approach is most effective. In addition, the organization of assessment has implications for the availability and location of data to support ecological valuation. It is important that data that are housed within individual program offices be made public and readily shared with other offices.

The *EBASP* contains suggestions for addressing some of the limitations on ecological valuation resulting from the Agency's internal structure. It advocates the creation of a high-level Agency oversight committee and a staff-level ecological benefits assessment forum. The committee endorses these efforts. (KS: Do we?? I added this but it has not been discussed by the committee.) Nonetheless, the Agency will continue to face significant external constraints when conducting ecological valuation. The committee recognizes the practical importance of these constraints and urges the Agency to be as comprehensive as possible in its analyses within the limitations imposed by these constraints. (KS: Should we say more/something different??)

3.3. An Illustrative Example of Ecosystem Benefit Assessment at EPA

In an effort to better understand the current state of ecosystem valuation at EPA, the committee examined in detail one specific case where benefit assessment was undertaken, namely, the Environmental and Economic Benefits Analysis that EPA prepared in support of new regulations for Concentrated Animal Feeding Operations (CAFOs) (U.S. Environmental Protection Agency 2002). [14,15] The Agency indicated that this analysis was illustrative of other EPA regulatory analyses of ecological benefits in form and general content.

Because the proposed new CAFO rule constituted a "significant regulatory action" under Executive Order 12866, EPA was required to assess the costs and benefits of the rule. [16] EPA identified a wide variety of potential "use" and "non-use" benefits as part of its analysis. [17] Using various economic valuation methods, EPA provided monetary quantifications in its CAFO report for seven environmental benefits.[18] Approximately eighty-five percent of the monetary benefits quantified by EPA were attributed to recreational use and non-use of affected waterways. According to Agency staff, EPA's analysis was driven by what it could monetize. EPA focused on those benefits for which data were known to be available for quantification of both the baseline condition and the likely changes from the proposed rule, and for translation of those changes into monetary equivalents. EPA's final benefits assessment provides only a brief discussion of the benefits that it could not monetize. The benefits table in the Executive Summary listed a variety of non-monetized benefits[19] but designated them only as "not monetized." EPA represented the aggregate effect of these "substantial additional environmental benefits" simply by attaching a "+B" place-holder to the estimated range of total monetized benefits. Although the Executive Summary gave a brief description of these "non-monetized" benefits, the remainder of the report devotes little attention to them.

Although considerable effort was invested in the CAFO benefits assessment, the assessment illustrates a number of limitations in the current state of ecosystem valuation at EPA. First, as noted above, in implementing the Executive Order, the CAFO analysis did not provide the full characterization of ecological benefits using quantitative and

qualitative information, as required by the OMB Circular A-4. Instead, the report focused on a limited set of environmental benefits, driven primarily by the ability to monetize these benefits using generally accepted models and existing value measures (benefit transfer).[20] These benefits did not include all of the major ecological benefits that the new CAFO rule would likely generate, nor all of the benefits that generated public support for the new rule. [21] The Circular requires that a benefit assessment identify and characterize all of the important benefits of the proposed rule, not simply those that can be monetized. By focusing only on a narrow set of benefits, the CAFO analysis and report understates the benefits of the rule change and distorts the rationale supporting the final rule. An unfortunate effect of this presentation is to suggest to readers that the monetized benefits constitute the principal justification for the CAFO rule.[22] Although in this case the focus on monetized benefits did not affect the outcome of the regulatory review, it is certainly possible that in a different context a benefits assessment based only on easily monetized benefits could inadvertently undermine support for a rule that would be justified based on a more inclusive characterization of benefits.

Second, the monetary values for many of the emphasized benefits were estimated through highly leveraged benefit transfers that were generally based on dated studies conducted in contexts quite different from the CAFO rule application.[23] This was undoubtedly driven to a large extent by time, data, and resource constraints, which make it very difficult for the Agency to conduct new surveys or studies and virtually force the Agency to monetize benefits using existing value estimates. However, reliance on dated studies in quite different contexts raises questions about the credibility or validity of the monetary benefit estimates. This is particularly true when values are presented as point estimates, without adequate recognition of the underlying limitations due to uncertainty and data quality.

Third, EPA apparently did not engage in a sufficiently comprehensive effort at the outset to model the rule's ecological impacts. The report presents only a simple conceptual model that traces outputs (a list of pollutants in manure – Exhibit 2-2 in the CAFO report) through pathways (Exhibit 2-1) to environmental and human health

effects.[24] This model provided useful guidance, but was not sufficiently comprehensive to assure thorough analysis of the rule's ecological impacts. As a consequence the analysis was unduly directed by Agency presumptions (or discoveries) about the availability of relevant data and the likely opportunities to quantify effects precisely and to link and monetize associated benefits. This was undoubtedly driven in part by the time pressures of putting together the regulatory impact analysis. However, without a comprehensive modeling effort at the outset, EPA had insufficient insight into the potential benefits that needed to be analyzed and valued. Developing integrated models of relevant ecosystems at the outset of a valuation project would also help in identifying important secondary effects, which frequently may be of even greater consequence or value than the primary effects.[25]

Fourth, the CAFO analysis clearly demonstrates the challenges of conducting ecological benefit assessments at the national level.[26] National rule-makings inevitably require EPA to generalize away from geographic specifics, both in terms of ecological impacts and associated values. However, it is possible (and desirable) to make use of intensive case studies (e.g., individual watersheds, lakes, streams, estuaries) in support of the national-scale analyses. Existing and ongoing research at local and regional scales offers more detailed data and models that could be better exploited, both to fill in gaps and to validate the national-scale analyses systematically. Systematically performing and documenting comparisons to intensive study sites could indicate the extent to which the national model needs to be adjusted for local/regional conditions and could provide data for estimating the range of error and uncertainty in the projected national-scale effects.

Fifth, although EPA invited public comment on the draft CAFO analysis as required by Executive Order 12866, there is no indication in the draft CAFO report that EPA consulted with the public during its analysis to help it identify, assess, and prioritize the effects and values addressed in its analysis, nor is there discussion in the final CAFO analysis of any comments received on the draft CAFO analysis. Early public involvement could play a valuable role in helping the Agency both a) identify all of the systems and services impacted by the proposed regulations and b) determine the

regulatory effects that are likely to be of greatest value. This would ensure that the benefits assessment includes the most important impacts.

Sixth, while EPA in its analysis and report appropriately emphasized the importance of using outside peer-reviewed data, methods, and models, EPA did not seek to peer review its application of them or its integration of these components in deriving benefit values for the CAFO rule. Once again, this is undoubtedly due in part to time and resource constraints. However, peer review, especially early in the process, would help EPA staff identify relevant and available data, models, and methods to support its analysis, and provide encouragement, direction, and sanction for more vigorous and effective pursuit of ecological and human wellbeing effects associated with the proposed rule. The general idea is to have individual components of the analysis (e.g., watershed modeling, air dispersal, human health, recreation, aesthetics) each reviewed, as well as a more general review of the overall analytic scheme.

Finally, EPA's analysis and report focused nearly exclusively on meeting the requirements as described in Executive Order 12866. This may not be surprising since the Executive Order provided the proximate reason for preparing the analysis and report. However, when EPA prepares a benefit assessment specifically to comply with Executive Order 12866, the Agency need not limit itself to the goals and requirements of the Executive Order. The Executive Order does not preclude EPA from adopting broader goals. The Executive Order provides merely that EPA shall conduct an "analysis" and "assessment" of the "benefits anticipated from the regulatory action" and, "to the extent feasible, a quantification of those benefits." By adopting a narrow focus, the report failed to consider or reflect the broader purposes that a benefit assessment can serve. Environmental benefit assessments such as the CAFO study can serve a variety of important purposes, including helping to educate policy-makers and the public more generally about the benefits that stem from EPA regulations, and it is important for EPA to recognize and have an incentive to consider this broader purpose.

4. AN INTEGRATED AND EXPANDED APPROACH TO ECOSYSTEM VALUATION: KEY FEATURES

The CAFO example discussed above highlights a number of limitations to the current state of ecosystem valuation at EPA. The committee's analysis points to the need for a comprehensive, integrated approach to valuing the ecological impacts of EPA actions, one that focuses on the impacts of most concern to people and integrates ecological analysis with valuation. This section describes an approach to ecological valuation developed and endorsed by the committee, followed in Section 5 by an overview of implementation. A more detailed discussion of the use of the approach and the methods that could be used to implement it are discussed in subsequent parts of the report (Parts 2 and 3). The approach should serve as a guide to EPA staff as they conduct RIAs and seek to implement the provisions of Circular A-4, as well as in decisions regarding regional and local priorities and activities.

The proposed approach has three key features, which are interrelated: a) a focus on impacts of most concern to people; b) an integration of ecological analysis and valuation; and c) inclusion of an expanded set of possible valuation methods.

The first feature reflects the committee's view that ecological valuation or benefit assessment should focus on the impacts or benefits that are likely to be most significant or of greatest concern to people, which might or might not be those that are most easily measured and monetized. This requires a systematic consideration of the many possible sources of value from ecosystem protection and an identification of the types of values that are providing the impetus for a particular policy change. Information about the ecosystem services or characteristics that are of greatest concern needs to be obtained early in the valuation process so that efforts at quantification and characterization of values can be focused on the related ecological changes. This requires a mapping of the effects of a given policy change on ecological assessment endpoints to the corresponding effects on ecosystems and ecological services. In addition, this focus will likely lead to an expansion of the types of services to be characterized, quantified, or explicitly valued. For example, even in the context of national rule-making, the inability to monetize a

specific benefit should not preclude its inclusion; if there is evidence that it is important to people, it should be included as a key component of total benefits and a detailed and careful (even if not monetized) characterization of that benefit should be provided.

The second key feature of the framework is the integration of ecological analysis with valuation. This integration needs to occur both at an early stage (in the identification of the impacts that matter) and at a later stage (when estimating the value of impacts). Thus, instead of having ecologists work independently initially to estimate ecological impacts and then "pass the baton" on to economists or other social scientists to value those impacts, the approach envisions collaborative work across disciplines to ensure that the analysis focuses on the impacts that are of greatest concern to society and that the ways in which these impacts are defined and measured are informative for valuation. This requires a committed interaction among the relevant bio-physical, ecological, and social scientists and analysts. The various disciplines must reach out to establish useful and credible links to each other. This interaction should commence at the beginning of the process and continue until the completion of the analysis. Ecological models need to be developed, modified, or extended to provide usable inputs for value assessments. Likewise, valuation methods and models need to be developed, modified, or extended to address important ecological/bio-physical effects that are currently underrepresented in value assessments.

Third, the approach draws on a variety of methods to characterize and measure the importance of changes in ecosystems, including economic methods, social/psychological assessments, and other methods based on bio-physical rankings or public or group expressions of value. It recognizes that different methods provide different ways of characterizing or providing information about values. Different methods could be used at different stages of the valuation process. For example, some methods might be well-suited to providing information that would be used early in the process to guide decisions about which ecological changes are likely to be most important to people, while other methods would be well-suited to quantifying or monetizing benefits that are specific to the EPA action. In addition, the suite of methods used could vary with the specific policy context, due to differences across contexts in: a)

information needs; b) the underlying sources of value being captured; c) data availability; and d) methodological limitations.

5. IMPLEMENTING THE INTEGRATED AND EXPANDED APPROACH

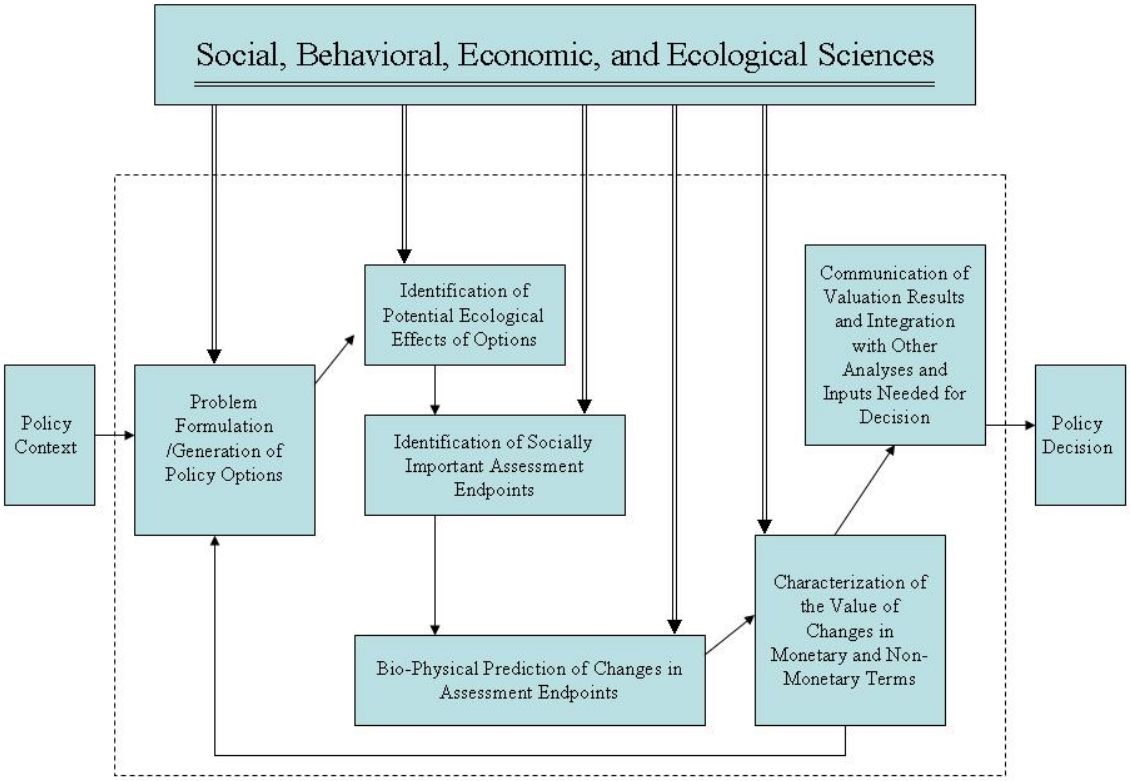
The previous section provides an overview of an integrated and expanded approach to ecological valuation proposed by the committee. This section discusses implementation of the approach in general terms. Part 2 of the report is devoted to a more detailed discussion of implementation in specific context, drawing on actual examples of decision contexts where ecological valuation could have played an important role.

The process for implementing the proposed framework would involve the following steps, depicted in Figure 2:

1. formulating the valuation problem and choosing policy options to be considered, given the policy context;
2. identifying the significant ecological changes that could result under the different options;
3. identifying the ecological changes that are socially important, i.e., choosing the assessment endpoints;
4. predicting the changes in the assessment endpoints in biophysical terms;
5. characterizing, representing, or measuring the value of changes in the assessment endpoints in monetary or non-monetary terms; and
6. communicating results to policymakers for use in policy decisions.

Although Figure 2 depicts these steps as sequential, in practice it is likely that iteration will occur at some points of the process. For example, information about the value of changes in assessment endpoints stemming from a given set of policy options might cause a reformulation of the problem or identification of alternative policy options that could be considered.

Figure 2: Process for Implementing an Expanded and Integrated Approach to Ecological Valuation



As depicted in Figure 2, the implementation of the approached in contingent upon the specific policy context. As noted above, ecological valuation can play a key role in a number of different decision contexts, including national rule-making and regional or local decisions regarding priorities and actions. The valuation problem should be formulated within the specific EPA context. Different contexts will generally be governed by different laws, principles, mandates, and public concerns. These contexts can differ not only in the required scale for the analysis (e.g., national vs. local) but possibly also in the type of valuation information that is needed. For example, in contexts where a benefit cost analysis is required, benefits need to be monetized whenever possible. In contrast, expressing benefits in monetary terms might be of little or no relevance to EPA analysts in other contexts, for example, when decisions are based on other criteria. Therefore the policy context in which the assessment is cast is a key influence on the appropriateness of data, models and methods.

Figure 2 also highlights the need for information and input from a wide range of disciplines at each step of the process, beginning with problem formulation. In addition, it suggests a structure that in many ways is parallel to the Agency's Framework for Ecological Risk Assessment (U.S. Environmental Protection Agency Risk Assessment Forum 1992). However, ecological valuation goes beyond the risk assessment framework by assessing the human values associated with predicted effects and hence provides a broader framework for assessment of ecological effects of EPA actions.

Several issues or considerations arise in implementing the steps of the process outlined above. A brief overview of these issues is provided here, as a prelude to the more detailed discussions in particular policy contexts that is included in Part 2.

5.1. Early Consideration of Effects that are Socially Important

A key component of the proposed approach is the identification and predictions of ecological changes that are important to people (steps 2 through 4). These could include both changes in the ecosystem itself that people value directly, or the resulting changes in the ecological services provided by those systems. The importance of a given change will depend on both the magnitude and bio-physical importance of the effect and the resulting importance to society.

Although Figure 2 suggests a linear process, this part of the process will generally be somewhat iterative. The first step is to determine a preliminary list of potentially important ecological effects, based on both the magnitude and bio-physical importance of the effect. Development of this list would draw primarily on ecological science. However, it is important to identify early in the process what effects people are likely to be concerned about. Consideration of what seems to be important to people can lead to a subsequent refinement of the list of ecological effects that will be the focus of any valuation. For example, are individuals likely to care about and hence value the re-introduction of native grasses into a marshland, or would they be just as happy with non-native grasses that perform similar ecological functions and have similar aesthetic appeal? Is animal waste disposal a concern to society primarily because of the recreational opportunities lost due to the resulting deterioration in water quality or is

society primarily concerned about other impacts? The range of ecological changes that are the focus of the valuation study needs to include the changes people care most about. Previous benefit assessments have often focused on what can be measured relatively easily rather than what is most important to society. This diminishes the relevance, usefulness and impact of the assessment.

An obvious question is how to assess the likely importance of different ecological impacts prior to completion of the valuation process. In fact, a main purpose of conducting a thorough valuation study is to provide an assessment of this importance. Nonetheless, in the early stages of the process, preliminary indicators of likely importance can be used as screening devices to provide guidance on the types of impacts that are likely to be of greatest concern. Relevant information can be obtained in a variety of ways. Examples range from in-depth studies of people's mental models and how their preferences are shaped by their conceptualization of ecosystems and ecological services, to more standard survey responses from prior or purpose-specific studies. In addition, early public involvement[27] or use of focus groups or workshops comprised of representative individuals from the affected population and relevant scientific experts can help to identify relevant or potentially important ecological changes for the specific context of interest.

In eliciting information about what matters to people, it is important to bear in mind that people's preferences depend on their mental models (i.e., their understandings of causal processes and relations), the information that is at hand to influence their understanding, and how that influence occurs. Expressions of what is important (e.g., in surveys) or of the tradeoffs people are willing to make can change with the amount and kind of information provided, as well as how it is provided. Collaborative interaction between analysts and public representatives can ensure that respondents have sufficient information when expressing views and preferences.

5.2. Challenges in Predicting Ecological Changes

The bio-physical impacts of a given EPA action can be identified at different temporal, spatial, and ecological levels. The latter include the individual level, the population level, the community level, the ecosystem level (union of biological

populations with their surrounding physical environment), and the level of the global biosphere. Living organisms supply goods and services that differ across all levels of organization, from the individual to the ecosystem or global biosphere. For example, the service provided by an individual animal unit is different from the service provided by a given animal population.

Estimating bio-physical impacts requires information about relevant ecological production functions. These functions provide a basis for estimation of the ecological changes that could result from a given EPA action or policy (e.g., percent reductions/avoidances of pollution in streams and lakes, reduced/avoided eutrophication of estuaries, reduced risk from the introduction of hormones and antibiotics into aquatic systems, improved/protected quality of community drinking water sources). In identifying and predicting ecological changes, it is important to consider their full range, including both primary and secondary effects, adequately accounting for uncertainty, stability of the system (including the effect of random shocks and management errors and the system's resilience), heterogeneity within a population or ecosystem, heterogeneity across populations or ecosystems, and dynamic changes in the ecosystem over time.

Numerous mathematical models of ecological production have been developed. These models cover the spectrum of biological organization and ecological hierarchy. Some have been developed for specific contexts (species, geographic locations, etc.) while others are more general. Primers on ecological theory and modeling (e.g., Roughgarden 1998) can provide a starting point for identifying available models.

Many of these ecological models have been developed to satisfy research objectives and not EPA policy or regulatory objectives. This poses at least three key challenges when using these models to assess the ecological benefits of EPA actions. The first challenge is to link existing models with Agency actions that are intended to control chemical, physical and biological sources of stress. The valuation framework outlined above requires an estimation of the bio-physical impacts that would stem from a specific EPA action. To be used for this purpose, ecological models must be linked to information about stressors. This link is often not a key feature of ecological models developed for research purposes.

Second, the outputs from the ecological models need to be linked to types of ecological services that are obtained from ecosystems and their structural and functional components. If adapted properly, ecological models can connect material outputs to stocks and services flows, assuming that the services have been well-identified. Providing the link between material outputs and services involves several steps, including identifying "service providers," determining the aspects of community structure that influence function, assessing the key environmental factors that influence the provision of services, and measuring the spatial and temporal scales over which services are provided (Kremen, 2005). Adaptation of existing models should consider alignment between: a) models; b) ecological services; c) ecological service providers; d) potential injuries; and e) the stressors under EPA purview.

Finally, the ecological models need to be appropriately parameterized for use in policy analysis. Numerous detailed ecological studies have been conducted at various levels, at Long-Term Ecological Research Sites, for example (Farber, Robert Costanza et al. 2006), which could provide a starting point for parameterizing policy-driven models. A key challenge is to determine whether (or to what extent) parameters estimated from a given study site or population at a given point in time can be "transferred" for use in evaluating ecological changes in a different location or time or at a different scale. In other words, to what extent are estimated parameters adaptable to the context of interest in estimating the benefits and values associated with EPA actions? In many cases, data do not currently exist to parameterize existing models so they can be used in assessing EPA's actions. Such data may need to be developed before the Agency can use these models fully. To the extent that transferable models and parameter estimates exist, it would be extremely valuable to have a central depository that EPA could draw on for this information.

Despite the challenges described above, the many ecological models that now exist provide a rich basis for developing particular models specialized for the various rule-making and other administrative actions EPA must consider. The EPA should, as a matter or ordinary practice, plan to formulate mathematical models of the systems to

which its actions pertain. It needs to acquire the expertise, both in house and through use of consultants, to carry out this part of its mission.

5.3. Approaches for Characterizing Values

Given predicted ecological changes, the value of these changes needs to be characterized and, when possible, measured or quantified. This requires that the output from the ecological impact assessment be in a form that can be used as an input in estimating the value of the change in ecosystems and/or ecosystem services. As noted previously, the associated value can stem either from a change in the flow of ecosystem services or a change in the ecosystem itself that is valued independently of the services it provides. Valuing the change in service flows involves two steps. The first is to translate bio-physical impacts into impacts on humans. The second is to value those human impacts. For values that are independent of service flows, the ecological changes can be valued directly.

To translate bio-physical impacts into human impacts, it is necessary to project how ecosystem changes will affect humans and the things that humans value. A variety of metrics exist that can provide quantitative information about some human impacts. These include the number and characteristics of the people/communities affected, social expressions of importance, the likely symptoms avoided or reduced, and the duration of the impact. There are at least three ways in which these metrics can provide very useful information. First, in some cases, these metrics may be used directly in policy decisions. For example, decisions based on human impact criteria (e.g., protection of children's health) may look directly to these measures as indicators of the appropriate policy choice. Second, they might be used as a proxy for some component of the benefits of ecosystem protection when that component cannot be readily valued. For example, in contexts requiring benefit cost analyses, the OMB Circular A-4 requires that benefits that cannot be monetized be quantified to the extent possible, and these metrics provide potentially useful forms of quantification. Finally, even when human impacts can be valued, these metrics provide information about human impacts that would presumably be relevant in the determination of the associated value of the ecological change. Thus, in all of these

contexts, estimates of the impact of the ecosystem change on human populations are needed.

In contexts where monetary metrics are required or desired and the necessary data and methods exist, the impact of the ecological change on the provision of some services to human populations may be translated into a monetary equivalent of that change using standard economic valuation techniques. For some valuation contexts economic methods for valuing changes are relatively well-developed. As noted previously, to date EPA ecological valuation efforts, such as the *EBASP* and the Star Grant program, have focused on valuing changes using economic methods. These methods are designed to estimate the benefit or cost of a given ecological change using a willingness-to-pay or willingness-to-accept measure of the utility equivalent of that change. They have been applied to the valuation of ecosystem services in a number of studies that have produced results that are useful for policy evaluation.

However, as in the CAFO study, economic valuation methods have generally been applied to a relatively narrow set of services. In some cases, these might not have been the services that people are most concerned about protecting. While there are continuing discussions about the role of economic valuations in principle, as a practical matter it is unlikely that all of the important benefits (or costs) of a change in ecological conditions can now, or soon will, be sufficiently captured by economic valuation methods. For this reason, the *EBASP* calls for exploring "supplemental" approaches to valuation. The valuation approach proposed by this committee calls for a more prominent role to be played by alternative methods for characterizing values, both as a practical alternative when economic methods cannot fully capture benefits because of data or other knowledge-based limitations and as a means of capturing the components of value that are not fully reflected in value measures based solely on economic measures of willingness to pay or willingness to accept. Expanding the methods "toolbox" to include other scientifically-based assessment approaches that can be applied along with or in place of economic assessments, where appropriate, will allow EPA to more fully represent the benefits of ecosystems and services. Of course, this toolbox should include only methods that meet accepted scientific standards of precision and reliability, are

appropriately responsive to relevant changes in ecosystems/services, and are properly related conceptually and empirically to things people value. For all methods, appropriate application will depend on the underlying scientific basis as well as the specific policy context.

The committee evaluated a number of different methods for characterizing values (described in detail in Part 3). These include social/psychological methods, which have been successfully used to identify and to assess a wide range of values that people hold and that have been important considerations for environmental policy and decision making. Social/psychological methods bear close resemblances to economic methods, but they do not seek to attain a unidimensional monetary measure of benefit, allowing instead for multiple dimensions of value to be expressed and considered by decision makers. Other approaches include assessments based on voting and other group expressions of social/civic values, as well as assessment methods based on bio-physical rankings that are potentially less directly dependent on human preferences and value judgments (although these clearly enter indirectly).

An expanded toolbox of methods could allow EPA to capture more completely the full range of benefits stemming from ecosystem protection. In addition, where resources allow, use of multiple methods to characterize the same underlying value can in some cases provide evidence of conjoint validity and increase the confidence of policy/decision makers and the public. Of course, a significant risk of applying multiple assessment methods to an environmental decision problem, even when multiple valuation methods are permitted by law, is that in particular contexts the methods may suggest conflicting information about relative values. In this case, it would be essential to try to ascertain the source of the differences. In some cases, differences may be readily explained by differences in the application of methodologies (e.g., eliciting values from different population groups or samples) or study limitations (e.g., inappropriate application of techniques or interpretation of results), or simply the inherent uncertainty that exists in estimating values as a result of from data limitations, theory limitations, and randomness (see discussion in Part 2). In other cases the differences may reflect the fact that the alternative methods are capturing fundamentally different sources or components

of value. In any case, information about the extent to which the different assessment methods yield similar or different conclusions about the value of an ecological change would be an important input into a policy decision.

5.4. Communicating Results

Information regarding the value of ecological changes stemming from EPA actions will only be useful in improving decision-making if it is communicated effectively to policymakers and integrated with other information used in policy decisions. In addition to policymakers, information about the value of ecological changes is likely to be of interest to community members and scientists alike.

Communicating the value of protecting ecological systems and services requires conveying not only value information, but also information about the nature and state of the ecological systems and services to which they apply and the ecological processes involved. Information can be and is often conveyed using mapped ecological information, other visualizations including photographs and graphs, ecological indicators and narratives. Integrated models with a geospatial interface, such as those developed by Costanza (Costanza and Farber 1986; Costanza, Sklar et al. 1990; Costanza 1993; Bockstael, Costanza et al. 1995; Fitz, DeBellevue et al. 1996; Cowling and Costanza 1997; Higgins, Turpie et al. 1997; Costanza 2002; Binder 2003; Costanza and Voinov 2003; Costanza 2004) are another approach to depicting the state of ecological systems and services. The SAB has proposed a framework for reporting on the condition of ecological resources (EPA, 2003). The EPA's draft Report on the Environment (U.S Environmental Protection Agency 2002) and reports of the Regional Environmental Monitoring and Assessment Program (REMAP) illustrate a range of approaches that can be used.

6. SUMMARY

Ecosystems play a crucial role in supporting life as we know it. They provide a wide array of services that directly or indirectly support or enhance human populations. In addition, they can be valued in their own right, for non-anthropocentric reasons stemming from ethical, religious, cultural or biocentric principles. Part of EPA's broad mission to protect human health and the environment includes the protection of ecosystems.

Many EPA actions affect the state of ecosystems and the services derived from them. However, to date ecosystem impacts have received relatively limited consideration in EPA policy analysis. It is imperative that EPA improve its ability to value ecosystems and their services to ensure that ecological impacts are adequately considered in the evaluation of EPA actions at the national, regional and local levels.

To date, ecological valuation at EPA has focused primarily on a limited set of ecological benefits. This stems primarily from the difficulty of predicting the impact of EPA actions on ecological systems and the services derived from them and the difficulty of quantifying, measuring, or characterizing the resulting benefits. The perception that benefits need to be monetized in order to be carefully characterized restricts the range of ecological impacts that are typically considered in EPA analyses, particularly at the national level.

The committee views EPA's efforts to improve its ability to value ecological systems and services as very important and timely. As EPA continues these efforts, the committee encourages the Agency to move toward covering an expanded range of important ecological effects and human considerations using an integrated approach. Such an approach would:

- Involve from the beginning an interdisciplinary collaboration among physical/biological and social scientists;
- Solicit from the beginning input from the public or representatives of individuals affected by the ecological changes.

- Recognize the many sources of value derived from ecosystems, including both instrumental and intrinsic values;
- Highlight the concept of ecosystem services and provide a mapping from changes in ecological systems to changes in services using the concept of an ecological production function;
- Expand the range of ecological changes that are valued, focusing on those changes in ecosystems and their services that are likely to be of greatest concern to people; and
- Explore and expand the use of methods that can appropriately characterize or measure the value associated with these changes;

Through the use of an expanded and integrated valuation framework of this type, EPA can move toward greater recognition and consideration of the effects that its actions have on ecosystems and the services they provide. In addition, it will allow EPA to improve environmental decision-making at the national, regional and local levels and contribute to EPA's overall mission regarding ecosystem protection. The remainder of this report develops the ideas embodied in this approach through a more detailed look at how the approach could be applied (Part 2) and the methods that might be used in implementing it (Part 3).

PART 2: APPLYING THE APPROACH

1. INTRODUCTION

Part 1 of this report presented an overview of an integrated and expanded approach to valuing ecological changes that result from EPA actions or decisions. The approach was described in general terms. In this part of the report, we discuss implementation of the approach in greater detail. The purpose is to discuss in more detail a number of issues that arise in implementing the approach and demonstrate its application in different EPA decision contexts in which ecological valuation can contribute to improved policy analysis and decisions. As background for this, the committee examined a number of examples of specific valuation contexts and used these examples to inform its views about application of the proposed approach. The discussion throughout this part of the report reflects the general insights gained from examination of these source examples.

Part 2 begins by examining in more detail the implementation of the concept of ecosystem services specifically in the context of valuation. This discussion identifies a number of issues that arise in implementing this concept. Part 2 turns then to a discussion of predicting ecological impacts, which is a key step in implementing the proposed approach. This is followed by a brief overview of the expanded set of valuation methods that were examined by the committee. (These methods are discussed in more detail in Part 3 of this report.) This overview is intended to provide the reader with basic information regarding the different methods at a level sufficient to allow the reader to follow the context write-ups that follow. The context write-ups discuss application of the proposed valuation approach in three different contexts where ecological valuation can play an important role for EPA: a) national rulemaking, b) site-specific assessments, and c) regional partnerships. Finally, this Part discusses issues regarding the uncertainty associated with ecological valuation and the communication of valuation results to both policy makers and the public.

2. IMPLEMENTING THE CONCEPT OF ECOSYSTEM SERVICES

Part 1 of this report includes an overview of the concept of ecosystem services, which forms a key element in the CVPESS valuation approach. While the general idea of ecosystem services is fairly clear, operationalizing this concept in a specific valuation context is often complicated and requires very careful consideration. This section discusses some issues that arise in applying the concept of ecosystem services in a valuation context.

An important step in implementation of the CVPESS valuation approach is a listing of the relevant ecosystem services. Such a listing raises the question of what to "count" or include in the definition of services. As noted previously, the Millennium Assessment used a very broad definition of ecosystem services that included both indirect and direct contributions of ecosystems to human well-being (ref). An advantage of this broad definition is that it recognizes the many different ways in which ecosystems contribute to life as we know it, and hence can lead to greater appreciation of the service role that various classes of biota play in providing services. An analysis by Weslawski, et al. (2004) for example, indicated that the invertebrate fauna found in soils and sediments are important in remineralization, waste treatment, biological control, gas and climate regulation and erosion and sedimentation control, yet the general public had no understanding or appreciation of these services in his analysis. They do have an appreciation of the higher level services or what could be termed "end-point services," such a clean water and aesthetics, and of course foods that could be derived from the system.

For some valuation purposes, it might be preferable to adopt a definition of ecosystem services that is narrower than the definition proposed by the Millennium Assessment and focuses instead on these end-point services. For example, Boyd and Banzhaf (2006) proposes a definition that focuses on services as "end products of nature", i.e., "components of nature, *directly* enjoyed, consumed or used to yield human well-being" [emphasis added]. They stresses the need to distinguish between (ecological) *inputs* (or intermediate products) and (ecological) *outputs* (end or final

products), and include only outputs in the definition of services. Under this definition, ecosystem functions and processes, such as nutrient recycling, are not considered services; while they contribute to the production of ecological end products or outputs, they are not outputs themselves. Likewise, this definition would not include goods or services like recreation that are produced by combining ecological inputs or outputs with conventional inputs (such as labor, capital, or time). In addition, Boyd and Banzhaf advocate defining changes in ecosystem services in terms of standardized units or quantities, which requires that they be measurable in practice. Such an approach is consistent with the concept of "green accounting," which extends the principles embodied in measuring marketed products to the measurement and consideration of the production, or changes in the stock, of ecological or other environmental "products" (reference NRC report by Nordhaus).

An advantage of defining ecosystem services in terms of end products is that, by highlighting the important distinction between inputs and outputs, it avoids the potential for double-counting. For example, for an ecological change that increased pollination, it would be double-counting to value both the improved pollination process (as an improved ecosystem service) and the increased agricultural output that results from it (as a separate service). Similarly, for a habitat improvement that leads to an increase in a bird population, it would be double-counting to value both the increased habitat (as an improved ecosystem service) and the resulting increase in the bird population (as a separate service). Doing so would be akin to valuing both the parts that went into production of an automobile and the final product, the automobile, itself. In principle, one can value a final product *either* directly (output valuation) or indirectly as the sum of the derived value of the inputs (input valuation), but not both, since separately valuing both intermediate and final products leads to double counting. Thus, in identifying and listing the ecosystem services to be valued, it is important to identify mutually exclusive services and to distinguish functions or processes that are inputs into the production of another ecological service so as to avoid double counting of the value of both the intermediate service (the input) and the final service (the output).

A careful delineation of ecosystem services in a specific context requires not only a determination of those ecological outputs that contribute to human well-being but also a

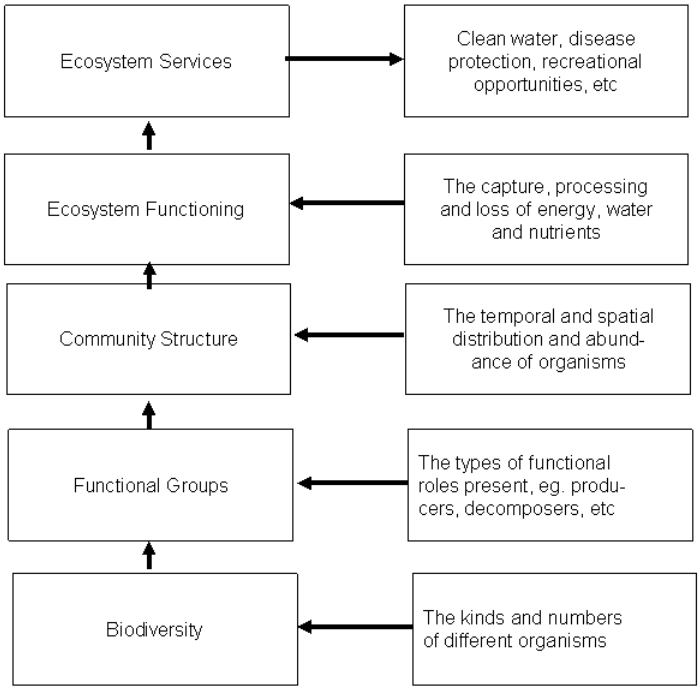
determination of the source of that contribution. This cannot be done deductively, but rather depends on what people value and why. For example, a list of ecosystem services by Wilson (2004) includes maintenance of riparian vegetation habitat for mammals, birds, amphibians, and insects. Even if the ecological role of this habitat is clear, the service that it provides to humans depends not only on its ecological role but also on what people care about. Do individuals care about the habitat per se (e.g., for aesthetic or moral reasons), or do they care about the populations it supports (e.g., the insect population)? If the value derives from the insect population, do individuals care about the insects for their own sake, or because they are a food source for fish that people care about? In the first case, the insect population itself provides a service (produced using habitat as an input) and should be valued directly. In the latter case, the insect population should not be treated as an end product or service; rather, the fish population provides the service (produced using the habitat, including the insect population, as an input), and we should value the change in fish resulting from the change in the insect population instead. Of course, a full delineation of services would also require information about why the fish population is valued, e.g., for its own sake or as an input into production of another good or service such as recreation or bald eagles.

Even with a clear delineation or listing of ecosystem services based on the concept of end products, ecological valuation requires an understanding of the functions, processes and components of the ecosystem that underlie and generate these services, which are inherently complex. Consider, for example, the ecological services associated with the activities of soil organisms that might be affected by disposal of waste on that soil. These organisms make their living from organic matter that is in, or added to, the soil. In the process of breaking it down for use certain groups maintain soil structure by their burrowing activities, which in turn provide pathways for the movement of water and air. Other kinds of organisms shred the organic material into smaller units that are in turn utilized by microbes that release nutrients in a form that can be utilized by higher plants for their growth, for example, or in dissolved form that enters into the water stream that leaves the immediate site into the water table or stream. Other groups of often-specialized microbes may release various nitrogen gases directly to the atmosphere. Thus, the nature of the soil organisms and the products that they utilize, store or release

regulates the biogeochemistry of the site as well as its hydrology and productivity and carbon storage capacity. These functions can be evaluated in general terms and related to the services that people more readily appreciate and value such as the capacity of the soil biota to process wastes and provide clean water (Wall, 2004).

Even when defined as end products, all of the links in the various levels of organizations of ecosystems are involved in the provisioning of ecosystem services, as indicated in Figure 3.

Figure 3: Illustration of How Abundance of Functional Groups Can Characterize the Complexity of an Ecosystem

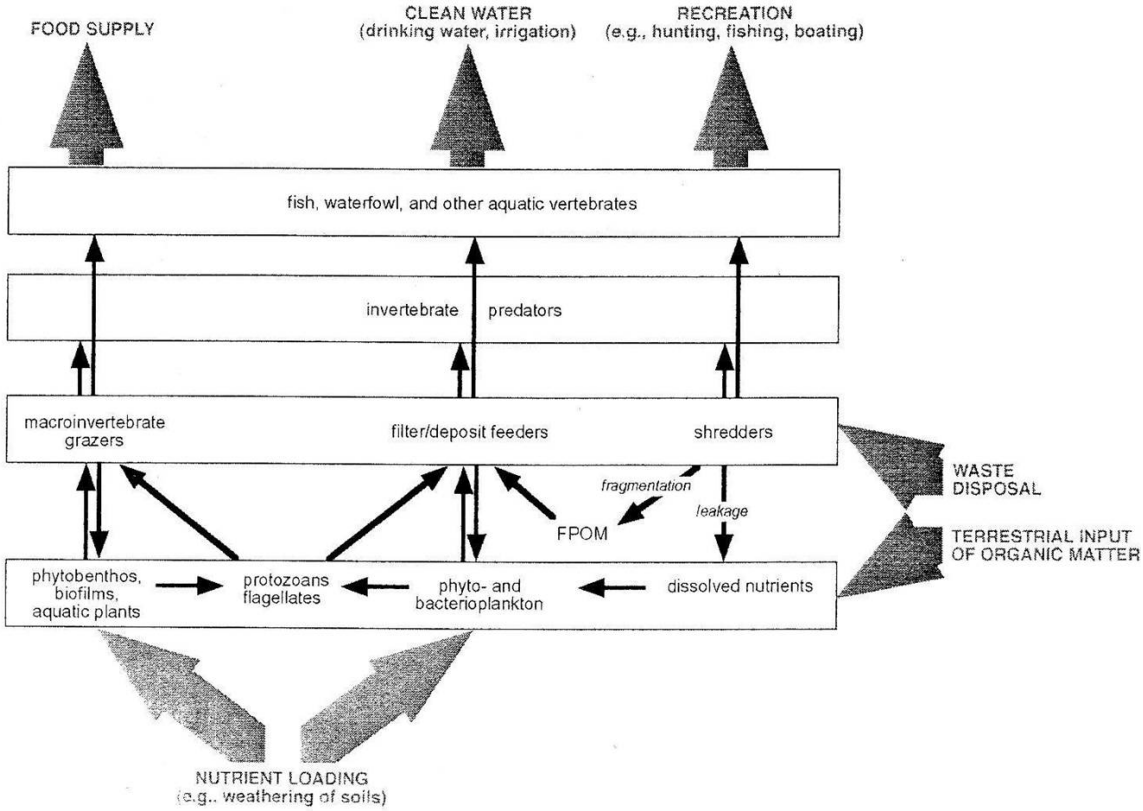


(Is the committee comfortable with use of this figure to illustrate the concept of functional groups? KS)

For a given ecosystem, the basic structure in Figure 3 can be used to develop a conceptual model of the functional levels of the ecosystem and how they contribute to the provision of ecosystem services. An example is illustrated in Figure 4 (taken from

Covich, et al., 2004). The important point of the figure is the emphasis on the major functional types and the role they play in providing the end point services. The key is identifying those components of each of the functional levels that are most directly related to the services of interest. Ecologists are at an early stage of linking ecosystem services with ecosystem functioning. Understanding these linkages better is an important research agenda.

Figure 4: Illustration from Covich et al., 2004, Showing Relationships of Major Functional Types to Ecological Services



Short of a full characterization of all of the linkages, it is possible to focus on groupings of organisms directly involved in the biological chain that affect the services of interest. This provides information about inputs as a proxy for the outputs. Because of the complexity of ecosystems they are often categorized not by species but by the abundance of the various functional groups present, e.g. decomposers of various kinds, as exemplified by the array of nematode types mentioned earlier, or bacteria that are specific

to producing a particular breakdown product, such as the methane producers, those that produce nitrous oxides, and so forth. The appeal of this approach is that within a given functional group there may be many different species that provide a given function even though one or more of the species of the group may not be present. In general, it is the functioning of a system that is of principal interest in terms of service provision, not "what species does the job." However, this principle does not hold at the recreational service level where a particular species, such as a given fish species, is the target of interest and the metric of concern.

The abundance of the groupings depicted in Figures 3 and 4 can be readily quantified. (Hal, can we give some references or examples? KS) There are readily available and fully tested techniques for evaluating all of the components in this chain. (Hal, can we give some references or be more specific? KS) For example, at the base of the ecosystem is its potential and realized biological diversity. Thus metrics that look at species richness and various diversity indices get at this directly. Through an analysis of the structures of the systems that are impacted, it should be possible to focus on functional types that are directly involved in providing the services of interest. For example, Weslawski et al. analyzed the services provided by various functional groups in estuaries and near-shelf sediment systems, providing a good starting point for relating functions to services. Ultimately, though, a better understanding is needed of how the various functional groups are affected by EPA actions and how these impacts in turn affect ecosystem services. Some taxonomic groups with wide functional diversity that are important in decomposition, such as the ubiquitous nematodes, offer promise in this regard for indicator purposes and have been so used in the past (Bongers and Ferris, 1999).

3. PREDICTION OF ECOLOGICAL EFFECTS

As noted above, the valuation process requires more than simply listing the relevant ecosystem components or services that might be affected by a given EPA action; it also requires a prediction of how those services or the associated ecosystems are or would be affected. To some extent, a preliminary prediction of impacts should already have played a role in the identification of relevant services, mainly as a screening device. However, after that initial screening, a more detailed and rigorous prediction of bio-physical effects is needed. To the extent possible, this prediction should be quantitative, using metrics that are either indicators of value or can be used as a direct input into an assessment of the associated value. In the context of national rulemaking, this quantification is necessary for values that will be monetized, and is required (to the extent possible) by the OMB Circular A-4 even for those values that cannot be readily monetized. In contexts where monetization is not required or desired, information about the magnitude of effects will still be a key component of value assessment.

3.1 The Road Map: A Conceptual Model

Formulation of a conceptual model is a key first step in predicting the ecological effects of EPA actions. This conceptual model should be constructed at a general level to guide the process rather than to incorporate more detailed analyses that will subsequently be considered in identifying the key interactions, assessing the endpoints and calculating the ecological valuations.

Development of the conceptual model is a significant task that deserves the attention of all the constituents of the process. These constituents include EPA staff from throughout the agency, experts in the relevant topics of consideration, and the public with a legitimate interest (standing) in the outcome. Although building the conceptual model is more complex than this linear sequence, in general the process involves these steps:

a) Building or importing a conceptual model of the ecological system(s), including preliminary decisions about the spatial and temporal boundaries. This step will be conducted largely by EPA staff and outside experts.

b) Inserting the likely stressors into the conceptual model; this may alter the initial and spatial boundaries. Most of this conceptualization will be accomplished by EPA staff and associated experts, but involving other constituents including the public at this stage will enhance transparency, provide the opportunity for more input and better understanding, and ultimately gives the process more legitimacy.

c) Identifying the significant ecological services requires the participation of all the constituencies as does anticipating ways in which the ecological services can be valued.

All three components (characterizing the ecological system, identifying the stressors, and identifying relevant endpoints and the significant ecological services) should be included in the beginning conceptual model. Building the conceptual model should be accomplished always with the recognition that one of the primary goals ultimately is to be able to value ecological services. This conceptual model, and the process for completing it and the embedded decisions within, should be a part of the formal record.

Characterizing the Relevant Ecosystem

Evaluating the ecological effects is an iterative rather than a linear process and EPA's process for ecological valuation should incorporate mechanisms for this iteration. For example, an action at a local site may initially be considered to affect only nearby regions. However, once the stressors are considered, the reactions may involve consequences in distant downstream watersheds or airsheds. Ecological effects may involve different persistence times (e.g., of carbon dioxide in the atmosphere), affecting both the temporal and spatial scales of the relevant ecological system. There are numerous studies, including EPA's regional analyses, risk analyses and the EMAP program, that provide guidance in identifying the proper boundaries and time scales for the ecological system under study as well as the ecosystem characteristics, stressors and endpoints (Harwell, et al, 1999, Young and Sanzone, 2002).

3.1. Identifying the Stressors

Although the identification of stressors to the relevant ecological system might be thought of as a relatively straightforward step, it too must be an iterative process. The important first step is a complete description of the project, whether it is a specific action on the ground or national rule promulgation. The description should include the quantity and quality of the anticipated stressors. As the stressors are identified in the context of the relevant ecological system, the conceptual model may need to be modified to incorporate additional stressors. For example, a relatively non-toxic chemical effluent might be considered insignificant as a stressor, but might become significant if the conceptual model indicated low stream flows or intermittent streams that would increase the concentration of the chemical to toxic levels during some parts of the year. Identifying socially important stressors and ecological endpoints can be accomplished by involving both experts and the public via several techniques, such as surveys, public meetings, focus groups, content analysis of public comments, solicitation of expert opinion and testimony, and summaries of previous decisions in similar circumstances.

3.2. Identifying Relevant Assessment Endpoints

Because of their inherent complexity, ecological systems cannot be characterized in their entirety, nor can their responses to stressors (endpoints) be completely measured and predicted. Because of the complexity of ecosystems, they are often categorized not by species but by the abundance of the various functional groups present as exemplified by nematode or bacteria types. There are a number of approaches to limiting the indicators to those that will provide the most direct information relevant to the services in question, for example, to focus on those functional groups that play a most prominent role in service (see previous discussion on ecosystem services).

An alternative approach to dealing with complexity is to rely on a set of generic assessment endpoints. EPA has developed a set of generic assessment endpoints that are based on environmental legislation and EPA's policies and precedents (Generic Ecological Assessment Endpoints, GEAE, 2003). These generic ecological assessment endpoints can be used as a starting point in the assessment of ecological effects. For example, Table 2 lists possible endpoints at the organism, population and community levels, including the policy relevance and the practicality of each endpoint.

Table 2: Table of Generic Assessment Endpoints Reproduced from U.S.EPA, 2003

Table 2-2. Generic ecological assessment endpoints (GEAEs): summary of the policy support for their use and their practicality[a]

GEAE #	Entity: attribute(s)	Policy support	Practicality
Organism-level endpoints			
1	Organisms: kills (mass mortality, conspicuous mortality)	Supported by many EPA programs; e.g., EPA has restricted the use of pesticides (e.g., diazinon and carbofuran) due to incidents of bird mortality.	Likelihood of kills from chemical pollutants can be estimated from toxicity testing. Incidents may be easy or difficult to observe, but when seen, they suggest a common mechanism or stressor exerting a strong effect.
2	Organisms: gross anomalies	Gross anomalies in birds, fish, shellfish, and other organisms are a cause for public concern and have been the basis for EPA regulatory action and guidance (e.g., assessed at Superfund sites, incorporated into biocriteria for water programs).	External gross anomalies are readily observed and are commonly included in survey protocols for fish and forests. They are also reported in toxicity tests of fish, birds, mammals, and plants.
3	Organisms: survival, fecundity, growth	Many EPA programs rely on organism-level attributes of survival, fecundity, and growth in assessing ecological risks (e.g., water quality criteria, pesticide and toxic chemical reviews, Superfund sites). Organism-level species protection is mandated by the Endangered Species Act, Marine Mammal Protection Act, Bald Eagle Protection Act, and Migratory Bird Treaty Act.	Results of toxicity tests of the survival, fecundity, and growth of organisms are abundant and often can be extrapolated to endangered species and other species of concern. Information on the ranges of listed endangered species is available through state and federal governments.
Population-level endpoints			
4	Assessment population: extirpation	EPA has taken action or provided guidance to prevent extirpation of local populations (e.g., assessment of likelihood of extirpation of fish populations due to acid rain). See also the description for Assessment population: abundance.	Extirpation can be predicted using population viability analysis. Demonstrating extirpation may be easy or difficult, depending on the conspicuousness of a species. See also the description for Assessment population: abundance.

Table 2-2. Generic ecological assessment endpoints (GEAEs): summary of the policy support for their use and their practicality[a] (continued)

GEAE #	Entity: attribute(s)	Policy support	Practicality
5	Assessment population: abundance	Major environmental statutes mandate protection of animals, plants, aquatic life, and living things generally, which can be inferred to entail protection of populations. EPA policies for pesticides, toxic chemicals, hazardous wastes, and air and water pollutants are intended to protect assessment populations of organisms. Mammals, birds, fish, aquatic invertebrates, and plants are typically assessed.	Changes in abundance may be predicted using conventional toxicity data with statistical extrapolation models and population models. OPPT evaluated a population model to explore effects of chloroparaffins on fish populations. Measurement of abundance in the field may be easy or difficult, depending on the species.
6	Assessment population: production	See description for Assessment population: abundance. Additionally, a number of laws are intended to maintain production of various economically valuable species. EPA water programs (e.g., National Estuary Program) and air programs (e.g., criteria pollutant standards) have involved protecting production of resource species populations.	Changes in production may be predicted using conventional toxicity data as well as population-based approaches. For resource species such as tree or fish species, production changes may be measurable in the field but may require long periods of observation.
Community and ecosystem-level endpoints			
7	Assessment communities, assemblages, and ecosystems: taxa richness	EPA water quality biocriteria frequently incorporate measures of community taxa richness. Additionally, EPA testing for pesticides, toxic chemicals, and water pollutants is intended to assess impacts to communities as well as populations and organisms. Fish, aquatic invertebrates, and aquatic plant assemblages are often assessed.	Changes in communities can be inferred or modeled from conventional toxicity data. Measuring taxa richness and abundance of aquatic communities, at least for fish and macroinvertebrate communities, is practical and well established. Ecosystem models that assess effects of toxicants on community properties are available and can use data acquired from organism-level laboratory testing, but they have not been routinely applied to date.
8	Assessment communities, assemblages, and ecosystems: abundance	As in the case of taxa richness, water quality biocriteria incorporate measures of community abundance, and EPA testing protocols are intended to assess impacts to communities.	See description above for taxa richness within assessment communities.

Tables such as this, expanded to include landscape-, regional and global-level endpoints (see EPA GEAE, 2003, Table 4.1; Harwell, et al. 1999; Young and Sanzone, 2002) can be used as a first step in characterizing the relevant ecological system and quantifying the responses to stressors—the ecological effects. Considerations prompted by the table can be helpful in constructing and evaluating the initial conceptual model. In identifying and predicting ecological changes, it is important to consider their full range, including both primary and secondary effects, adequately accounting for uncertainty,

stability of the system (including the effect of random shocks and management errors and the system's resilience), heterogeneity within a population or ecosystem, heterogeneity across populations or ecosystems, and dynamic changes in the ecosystem over time (see Part 1 of this Report).

3.3. The Use of Ecological Models

While a conceptual model can provide a road map for predicting ecological effects, specific ecological models are needed to quantify effects and incorporate of the dynamic interactions among the ecosystem components, such as interactions among species, changing dynamics of population numbers with alterations in habitats, or accumulation of toxic materials in substrates with different absorption capacities. Because of the complexity of most ecosystems, models are used to organize information, elicit the interactions among the variables represented in the models, and when run under different sets of assumptions or driving variables, to predict outcomes. [28] Models become imperative to determine aspects of ecosystem structure that will influence future service production. The choice of models, and the availability and appropriateness of supporting databases, will be different depending on the scale of analysis (e.g., local vs. national) and the precision of the question or hypothesis to be evaluated.

There are numerous ecological models that are used to describe ecological "systems" and various production functions, including scales from individual plants to regional characteristics such as crop productivity to continental migration of large animals. These models frequently focus on specific ecological characteristics, for example, populations of one or more species or the movement of nutrients through ecosystems. Models cover the spectrum of biological organization and ecological hierarchy. Primers on ecological theory and modeling such as Roughgarden 1998, Primer of Ecological Theory can provide a starting point for identifying available models. Some statistical models are relatively small, containing a few equations. Other ecological models are very large, involving hundreds of interacting calculations. Although many of these ecological models have been developed to satisfy research objectives and not agency policy or regulatory objectives, they are adaptable for use in estimating the benefits and values associated with agency actions. In fact, EPA currently employs a

number of ecological models, ranging from fairly straightforward toxicity models to population models to regional landscape models.

The primary focus of ecological models has been on understanding the dynamics in ecological systems, including for example, the role of abiotic driving variables on production, the interaction among species and the rate of carbon sequestration on continental scales. Although many of these models are well established and are used routinely for describing ecological systems, the results from all ecological models are approximations—they are estimates with known or unknown levels of statistical uncertainty—and no ecological model includes all the possible interactions. Because of the inherent complexity of ecological systems, most models are descriptive and predictions are more tentative, especially as ecological systems are frequently non-linear and subject to threshold changes. Although some ecological models explicitly or implicitly incorporate human dimensions, many of them focus primarily on ecological functions. Finally, the applicability, and to some degree the formulation of ecological models is frequently constrained by the insufficiency of data to build and test the models.

Because so many ecological models exist, because none of them explicitly represent all the possible variables and their interactions, and because none is proven to be completely "accurate" under all defined circumstances, EPA is faced with deciding which models to employ for evaluating potential changes in endpoints at the site, regional and national scales. In theory, EPA could outline the types of ecological conditions under which it expected to consider risks and impacts, inventory the existing ecological models, conduct an assessment of their effectiveness and then offer a catalog of "approved" models with specifications and restrictions for their application. Although such an approach would have some appeal, it does not accommodate the dynamics of the scientific process, namely that existing models are always being modified on the basis of new understanding or additional data. Moreover, new models are continually being created and tested. Such a catalog of approved models would have some utility in the sense that use of these models would imply a level of credibility and acceptability that would not otherwise need to be re-established with every new assessment. In addition, such a catalog of approved models would at some level create greater consistency among

the methods used in the various EPA regions, presumably evolving toward a smaller number of models with greater validity.

The alternative approach recognizes the dynamics of evolving science, and specifies prerequisite characteristics of models rather than specifying particular models. Under this approach, models would be selected in the judgment of EPA that best address the particular issue—so long as certain criteria were met. For example, EPA could mandate that models and data sets used to predict changes in ecological endpoints should meet the following seven conditions:

a) A beginning conceptual model that identifies, at least in a preliminary way, the state of the ecological system, the likely stressors and responses to those stressors and all the socially important anticipated interactions.

b) Utilization of databases that are in existence for the site, region or country that can provide, at a minimum a first approximation of the probable changes in endpoints. These more general data sets may need to be refined for the specific region or site depending on the project or the rule being considered, but initial assessments using these more generalized data sets will produce a range of likely outcomes which may be analyzed in more detail.

c) Adaptation of existing models should consider the congruent alignments among: (1) models; (2) ecological systems; (3) ecological services; (4) ecological service providers; (5) potential injuries; and (6) the stressors under EPA purview.

d) Models that are sufficiently comprehensive and have been used repeatedly so that there is a sufficient depth of understanding about their implicit assumptions, their reliability (robustness) and the reasonable range of applicability (space and time scales). These models would have been subjected to sensitivity analysis so there was a well-defined domain of outcomes from stochastic inputs.

e) Analytic output from the models should include a measure of variance that can be used to describe uncertainty in the predicted outcomes in a statistical distribution.

f) Results (production functions) from the analytical model should provide outcomes that are amenable to monetary and non-monetary valuation techniques.

g) Results from the models should provide guidance in a form that not only can be subjected to valuation techniques, but is readily usable by managers and rule- and policy-makers as well as by the interested public.

All of this emphasizes the importance of continued research aimed not only at improving understanding of ecological systems, but in particular at identifying the minimum information requirements for adequately describing and modeling the properties of ecological systems that result in important ecological services.

3.4. Gap between ecological models and the needs of ecological valuation

There is currently a gap between the outputs of most ecological models and the inputs required for valuation of ecological services. This gap arises for two general reasons. First, ecological models have largely focused on describing ecological systems in terms of ecological structure and function rather than in terms of social values. That is, the links between outputs of some ecological models and human uses of the ecosystem are not known or easily quantified. Many of these models offer powerful comparisons among ecosystems as they are intrinsically different or respond differently to stressors or changes in driving variables. As such, these model outputs may or may not be cast in terms of direct concern to people, and as such are not designed as inputs to valuation techniques. For example, evapotranspiration rates, rates of carbon turnover and changes in leaf area are important for ecological understanding, but are not easily translated in to values of direct human importance. Of course, there are some examples of models with outputs directly related to human values, such as those that predict fish and game populations or forest productivity. However, these represent a limited set of ecosystem services.

The second reason for the gap between ecological models and valuation needs relates to the complexity of ecological systems and their dependence on an array of site-specific driving variables. Because of this, many ecological models are site specific. Moreover, the relatively large amounts of site-specific data required to build and

parameterize models means that their transferability is limited, either because the model has been developed using spatially constrained data or because inadequate data are available at secondary sites with which to drive the model. This site-specificity may significantly limit the models' applicability to the spatial and temporal complexities required in valuing ecological services, certainly at regional and national scales. The absence of key data was identified in an analysis of the Millennium Assessment (Carpenter, et al. 2006): *serious constraints in the Millennium Assessment included systematic information on: stocks, flows, and economic values of many ecosystem services (e.g., freshwater fisheries, natural hazard regulation, groundwater, and pollination); knowledge of trends in human reliance on ecosystem services, particularly services without market value (e.g., domestic fuel wood and fodder; systematic local and regional assessments of the value of ecosystem services; and connections between data on human systems and ecosystems*.

Generic economic models and ecological models are useful, but in particular places (specific social values, specific ecological systems) either or both may fail to a degree because of the lack of specific information and the cost to obtain the information necessary for one or both models. There are numerous inter-twined challenges in the development of both economic and ecological models, including incorporating the required range of primary and secondary effects, characterizing the stability and dynamics of the systems, and dealing with the multiple-dimensioned heterogeneity of human populations and ecosystems.

3.5. Closing the Gap

Predictions of ecological effects of EPA actions are used for two fundamental purposes: as direct input into valuation methods (socio-psychological, economic, mediated modeling, etc.), and to quantify impacts when they cannot be monetized (in accordance with guidance in OMB Circular A-4) or when monetization is not necessary or desired. To date, direct measurements of ecological services and models that predict ecological services have been primarily aimed at describing and comparing the characteristics of ecological systems, not providing inputs to various methods of quantitative or qualitative valuations. However, there are now several promising

directions for closing the gap between ecological measurements and model outputs and valuation inputs. These advances are in three areas:

- Organization of ecological data sets
- Evaluation of benefits transfer of ecological data
- Advances in modeling ecological services

Organization of Ecological Data:. Data on the structure and function of ecological systems are becoming much more available and better organized across the country. Part of the increased availability is simply that web-based publication now enables authors to easily post data and further analysis in electronic forms, available to other researchers. Also, as governmental agencies are being held more accountable, data used in decision-making are expected to be made available to constituents.

Within the ecological research community, the National Science Foundation (NSF) Long-Term Ecological Research (LTER) program has had an emphasis on organizing and sharing data in easily accessible electronic datasets. Although these data were rarely collected for the purpose of valuing ecological services, they are particularly valuable because they frequently measure long-term trends. As such, these data are useful in separating short-term fluctuations from longer term patterns in ecological properties. Also, the LTER program more recently has focused on "regionalization" in which data from sites surrounding the primary site are collected, thus providing a regional context for site-based measurements and models. Planning for the forthcoming NSF National Ecological Observatory Network (NEON) includes a Networking Information and Baseline Design (NIBD) component, which connects the key scientific questions to the data required to answer the questions. It will be important for EPA to have effective links into the NEON planning process, and to expand its involvement with the NSF LTER program, which is now undergoing a major refreshing of its research and data sharing protocols.

Evaluation of benefits transfer of ecological data. Despite the increasing availability and organization of ecological data, the costs are too prohibitive to allow

extensive data to be collected from all the sites on which EPA is considering action. From an ecological perspective, therefore, the issue is the reliability of transferring ecological information from one site or study to other sites or over different spatial or temporal scales. Information in this sense can include tools or approaches, data on properties of an ecosystem or its components, and services or benefits derived from an ecosystem.

There are no hard and fast rules for when ecological information can be transferred, and the confidence with which information can be transferred depends on the type of information and the system in question. Given the complexity, richness of interactions and the propensity for non-linearity, extrapolation of ecological information requires caution. However, certain generalizations are possible. Information is more likely to be transferable with greater similarity between and among contexts, e.g., ecosystems. Also, aggregate information, such as data on ecosystem properties are more likely to be transferable than information on particular species or the interactions of particular species. Thus, the ecosystem properties (e.g., leaf area index, primary productivity, nitrogen cycling patterns) of an oak-hickory deciduous forest in Tennessee might be transferable to oak-hickory forests in other parts of the eastern U.S. that are at similar stages of development. To a lesser extent, the information might be transferable to other types of deciduous forests.

Information could be transferable to other spatial or temporal scales if the dynamics over time and space scales are known for the ecosystem. For instance, if data are available on how the characteristics of an oak-hickory forest change as it develops or through cycles of disturbance, then it should be possible to transfer data from one point in time to another. Similarly, if information is available on how the properties of the system vary with spatial environmental variation (local climate, soil type, land-use history), then it should be possible to extend information from one spatial context to another. EPA and other national and international agencies have sponsored extensive research on “scaling up” of data from particular sites to regions, and results from these analyses are applicable to the transfer of information on ecological properties and services.

To some extent, the same generalizations apply to transferring tools such as models, although success depends on how generally applicable the tool is and how

difficult (in terms of data requirements) it is to parameterize for other situations. For example, forest ecosystem models can often be transferred to other forests using available information from sources such as LTER sites.

The applicability of transferring benefits depends on characteristics of related resources and conditions, and also, on the reasonableness of using a static definition of an economic trade-off in a dynamic ecological system. Farber, et al. (2006) have attempted to classify the benefits transfer of ecosystem services from one context to another (see Table 3 below). In some cases, e.g., carbon sequestration (gas regulation) the transfer is appropriate at large spatial scales; in other cases, the processes operate at small scales but the processes are so general that they can be transferred with high confidence (e.g., value of game harvest). Some characteristics, such as genetic biodiversity (genetic resources) or spiritual values are very site-specific and thus the benefits cannot be transferred with confidence.

Table 3: Farber et al., 2006), Classification of Benefits Transfer of Ecosystem Services from One Context to Another

Ecological Service	Transferability
Gas regulation	High
Climate regulation	High
Disturbance regulation	Medium
Biological regulation	High
Water regulation	Medium
Soil retention	Medium
Waste regulation	Medium/high
Nutrient regulation	Medium
Water supply	Medium
Food	High
Raw materials	High
Genetic resources	Low
Medicinal resources	High
Ornamental resources	Medium
Recreation	Low

Aesthetics	Low
Science and education	High
Spiritual and historical	Low

Advances in modeling ecological services. Over the past two decades, appreciation for the role of ecological services has increased among ecological researchers and decision-makers, and much more slowly among members of the public—those affected by EPA rule-making. As an example, ecological researchers know that invertebrates in soils and sediments are important in re-mineralization, waste treatment, biological control, gas and climate regulation and erosion and sedimentation control yet Weslawski, et al (2004) found that the general public had no understanding or appreciation of these services. On the other hand, there are numerous ecological models with outputs the public does understand, such as those that predict wildlife populations for hunting permits, permissible fish catches and algae levels in relation to fertilizers in lakes and streams. The challenge is to predict changes in a broader set of ecological services in a form that is readily interpreted and used in a valuation context.

As noted above, EPA currently relies on a number of reasonably well-established models to predict ecological effects, recognizing that these models do not include all the possible considerations, especially in complicated issues or those that prevail over extended space and time scales. In some cases, the Agency has coupled these predicted effects with estimates of willingness to pay to generate monetized values of ecological changes. However, the application of this approach has been limited to a fairly small number of ecosystem services (see discussion of the CAFO example in Part 1 of this report). With the acquisition of increasing amounts of data, greater confidence in transferring data benefits and with greater model maturity derived from more experience, the process will allow for broader application of this approach. This approach, based on the use of existing ecological models, has the advantage that it fits many of the seven criteria listed above. In addition, it has been tested or used in a number of circumstances so its reliability can be assessed.

Although reliance on a single existing ecosystem model is one approach for predicting ecological effects of EPA actions, there are two other approaches that EPA

should consider. The first is based on the use of indicators (a form of "simplification") while the second is based on meta-analysis (a form of "data aggregation"). Each is briefly discussed below. (Paul: Is this terminology ok? It seems clearer to me than "simplification" and "data aggregation" approaches. KS)

The first approach involves selecting key predictive variables or indicators rather than attempting to measure and value all the possible significant outputs (e.g. Karr, 1993). Trends in ecosystem services are often most effectively communicated through indicators that simplify and synthesize underlying complexity. In addition, the use of large, complexity ecological models can be difficult pragmatically, especially because of the quantities of required data and the time to implement. As a result, making numerous or rapid evaluations is difficult (Hoagland and Jin, 2006) and simplification would be far more practical. Thus, the use of indicators can have advantages in terms of both generating and conveying information about ecological effects.

Many ecosystem indicators have been proposed (EPA/EC, 1996; National Research Council, 2000) and several states have sought to define a relatively small set of indicators of environmental quality to convey the value of ecological services. There currently is no agreement on a common set of indicators that can be consistently applied and serves the needs of decision makers and researchers in all contexts (Carpenter, et al., 2006). However, there are guidelines for specific issues. For example, in evaluating the economic consequences of species invasion, Leung, et al. (2005) have developed a framework for rapid assessments to guide in prevention and control, simplifying the ecological complexity to a relatively small number of easily estimated parameters. Because of the complexity of the interactions between economic and ecological systems, economists frequently take a similar simplification approach that focuses on effects occurring only in the relevant markets, assuming that the effects on the broader market are negligible and can be ignored (Settle, 2002).

This simplification approach to ecological modeling will never satisfy those who will always want to identify all the possible consequences of EPA actions. For example, Barbier's study of the economics of species invasion involved a predator-prey model with inter-specific competition and dispersion. The model results demonstrated that the extent to which commercial fishing was reduced by the introduction and spread of invasive

species was determined by the types of ecological interaction. He further argues that future models should consider more complex ecological interactions, habitat modification and non-market damages (Hoagland and Jin, 2006). The question, of course, is the practicality of building ever more complex models that must address a wide array of issues over multiple spatial and temporal scales. It may well be that with accumulated experience, the simplified model of selecting a few key indicators or ecological processes that can be valued may prove to be the most practical approach.

Similarly, there are ecological frameworks designed to incorporate multiple dimensions into a coherent presentation that describes the status of ecosystems within a region, especially as they relate to social values. For example, the "ecosystem report card" in South Florida (Harwell, et al., 1999) is based on particularly germane criteria:

- be understandable to multiple audiences,
- address differences in ecosystem responses across time, '
- show the status of the ecosystem
- characterize the selected endpoints, and
- transparently provide the scientific basis for the assigned grades on the report card.

The report card identifies seven essential ecosystem characteristics that are thought to be of importance, i.e., habitat quality, integrity of the biotic community, ecological processes, water quality, hydrological system, disturbance regime (changes from natural variability), and sediment/soil quality that were then related to the goals and objectives for the ecosystem integrity report card.[29] Related ecological endpoints were selected based on both scientific issues and societal values. The outputs are not designed to be monetized, but rather are described by narratives or quantitative/qualitative grades that are scientifically credible and easily understood by the public. (Note: There are other examples of using report cards to characterize the status of a given ecosystem. The extension of this idea, of course, is to use changes in the grades as indicators of ecological effects of EPA actions. Do we want to advocate the report card approach as a

way to "characterize" benefits for purposes of Circular A-4 when they can't readily be monetized? KS)

A third alternative, the use of meta-analysis, involves collecting data from multiple sources and attempting to draw out consistent patterns and relationships. For example, Worm, et al. (2006) attempted to measure the impacts of biodiversity loss on ecosystem services across the global oceans. They combined available data from multiple sources, ranging from small-scale experiments to global fisheries. In these analyses, it is impossible to separate correlation and causation, which is a severe limitation. On the other hand, by examining data from site-specific studies, coastal regional analyses and global catch databases, at least correlative relationships could be drawn between biodiversity and decreases in commercial fish populations—variables that can be monetized.

In a similar "data aggregation" approach, de Zwart, et al. (2006) noted that ecological methods for measuring the magnitude of biological degradation in aquatic communities are well established (e.g. Karr, 1981), but determining probable causes is usually left to a combination of expert opinion, multivariate statistics and weighing of evidence. As a result the results are difficult to interpret and communicate, particularly because mixtures of potentially toxic compounds are frequently part of these assessments. To address this issue the authors used a combination of ecological, ecotoxicological and exposure modeling to provide statistical estimates of probable effects of different natural and anthropogenic stressors to fish assemblages. This approach: a) links fish, habitat, and chemistry data collected from hundreds of sites in Ohio streams; b) assesses the biological condition at each site; c) attributes impairment [e.g., loss of one or more of 117 fish species] to multiple probable causes; and d) provides the results of the analyses in simple-to-interpret pie charts. When data were aggregated from throughout Ohio, 50% of the biological effect was associated with unknown factors and model error; the remaining 50% was associated with alteration in stream chemistry and habitat. While the results are not perfect, the point is that the technique combines multiple data sets and assessment tools (models) to arrive as estimates of loss of fish species based on broad patterns. Thus, like the previous study of the relationship of biodiversity to ocean productivity, this study

aggregates data from many sources and uses various models to arrive at estimates which can be easily interpreted and at least in the case of game fish species, can be monetized.

In summary, EPA can continue to refine the models it uses, paying particular attention to the six principles described above as a screen for this model selection process. The success of this approach will require continued expansion of data collection and increasingly demanding data management systems. In addition, EPA can explore the possibility of selecting key variables or indicators that are highly correlated with other ecological services. Finally, EPA can also focus on various levels of data aggregation that enable meta-analyses to identify broad relationships that obviate the need for ever more detailed data collection and model construction.

4. INTRODUCTION TO DIFFERENT TYPES OF METHODS

The process for implementing the CVPESS approach requires the use of an expanded set of methods for characterizing the value of the predicted ecological effects of EPA actions. In this section, we provide a brief overview of the categories of methods that the committee evaluated. The purpose is to provide the reader with enough information to follow the subsequent discussions about applying the approach in different policy contexts. A detailed discussion of specific methods is included in Part 3 of this report. In addition, Appendix A provides detailed information about survey methods.

4.1. Biophysical Ranking Methods

In some contexts, policymakers or analysts define values based on quantification of bio-physical indicators. Possible indicators include species biodiversity, bio-mass production, carbon sequestration or energy and materials use/redistribution/flows. Quantification of ecological changes in bio-physical terms allows these changes to be ranked based on individual or aggregate indicators for use in evaluating policy options. Use of a biophysical ranking does not explicitly incorporate varying human values based on human preferences. Rather, it reflects either an alternative theory of value (based, for example, on energy flows) or a presumption that the indicators provide a proxy for human value or social preferences. This latter presumption is predicated on the belief that the healthy functioning and sustainability of ecosystems is fundamentally important to the well-being of human societies, and all living things, and that the benefits of any change in ecosystems can be assessed in terms of the calculated effects on overall ecosystems health and sustainability. Some people view the fact that these ranking methods are not directly tied to human preferences as a drawback, while others view it as a positive feature of these methods.

The committee evaluated two types of bio-physical rankings. The first was a ranking method based on conservation value. This method develops a spatially-differentiated index of conservation value across a landscape based on an assessment of rarity, persistence, threat, and other landscape attributes, reflecting their contribution to sustained ecosystem diversity and integrity. The method provides a scientifically-based approach to assigning conservation values that can used by policymakers or stakeholders

to prioritize land for, for example, acquisition and conservation. Based on GIS technology, it has the capability to combine information about a variety of ecosystem characteristics and services across a given landscape, and to overlay ecological information with other spatial data. In addition, data layers can be used for multiple policy contexts. Conservation values have been used in various contexts by federal agencies (e.g., Forest Service, Fish and Wildlife, National Park Service, and Bureau of Land Management) as well as by non-governmental organizations (e.g., the Nature Conservancy, NatureServe) and regional and local planning agencies.

The second group of bio-physical methods that the committee evaluated were based on energy and material flows. Energy and material flow analysis is the quantification of the flows of energy and materials through complex ecological and/or economic systems. These analyses are based on an application of the first (conservation of mass and energy) and second (entropy) laws of thermodynamics to ecological-economic systems. Examples include embodied energy, emergy, and ecological footprints. Of these three, embodied energy and ecological footprints are based on a consistent set of principles, while emergy is not (and is hence not scientifically sound). Embodied energy measures the (available) energy cost of goods and services using input-output analysis or flow accounting methods. Ecological footprint analysis also uses input-output analysis, but measures "costs" in land units (rather than energy units) based on the biologically productive land area (rather than the amount of energy) required to meet various consumption patterns. These techniques have been used to estimate implicit costs or "shadow prices" of providing ecosystem goods and services, measured in physical rather than monetary units. While such costs can be used to rank alternatives based, for example, on an energy theory of value, they will provide a proxy for preference-based values only under limited conditions (see Part 3, section 8).

4.2. Social-Psychological Methods

Social/psychological methods seek to characterize the values that are held, expressed, and advocated by people. They focus on individuals' judgments of the relative importance of, acceptance of, or preferences for ecological changes. Individuals making the judgments may respond on their own behalf or on behalf of others (society at large or specified sub-groups) and the basis for judgments may be changes in individual

well-being, or civic or ethical/moral obligations relevant to ecosystems and ecosystem services. That is, people may hold, express and advocate bio-ecological values or ethical values that are unrelated or even counter to their own wants and needs.

Social/psychological methods provide scientific means for determining people's value-relevant perceptions and judgments about a wide array of objects, events and conditions. They typically focus on choices or ratings among sets of alternative policies, and may include comparisons with potentially competing social and economic goals.

Social-psychological methods elicit information about preferences and values primarily through surveys, focus groups, and individual narratives. However, recently experts in this field are also experimenting with eliciting this information through observations of behavioral responses by individuals interacting with either actual or computer simulated environments.

Survey questions are typically framed as choices (among two or more options), rankings, or ratings. Survey methods that may be especially important in the context of ecological valuation include perceptual surveys (e.g., assessment of ecosystem attributes) and conjoint survey methods (e.g., choice among different combinations of ecosystem attributes). Quantitative analysis of responses are usually interpreted as ordinal rankings or rough interval scale measures that provide relative measures of differences in assessed values. Similarities and differences among different segments of the public can also be identified and articulated. Surveys may be especially useful when the values at issue are difficult to express or conceive in monetary terms or where monetary expressions are viewed as ethically inappropriate. Surveys to elicit value-related information have been used extensively by other federal agencies, including the U.S. Department of Agriculture's Forest Service.

In contrast to surveys, which are based on large samples and are not interactive, focus group methods elicit information about values and preferences from small groups of relevant stakeholders engaging in group discussion lead by a facilitator. Collective narratives are subjected to qualitative analyses to identify and possibly to ascertain levels of consensus on relevant issues, perspectives, and positions represented by participants. Given the small number of participants, the goal of a focus group is rarely value assessment per se, but rather an articulation of all of the values that may be relevant. Use

of focus groups early in the decision process can help in identifying ecosystem effects that might be particularly important to the public. Individual narratives can also be informative in this way, although individual participants typically come from a more narrowly-defined target group.

4.3. Economic Methods

The economic approach to valuation is an anthropocentric approach based on utilitarian principles. It includes consideration of both instrumental values and intrinsic values, but only to the extent that preservation based on intrinsic value contributes to an individual's welfare. Because it is utilitarian-based, it assumes there is the potential for substitutability between the different sources of value that contribute to welfare. In addition, it assumes that individual preferences, which determine the degree of substitutability for that person, are well-formed. Most of EPA's work to date on ecological valuation has been based on the use of economic methods, and these methods are the focus of the recently released *EBASP*.

The concept of value underlying economic valuation methods is based on substitutability, or, more specifically, on the tradeoffs individuals are willing to make for ecological improvements or to avoid ecological degradation. By itself, an ecological change that an individual values will increase that person's utility. The value or benefit of that change is defined to be the amount of another good (typically money) that the individual is willing to give up to enjoy that change (willingness-to-pay) or the amount of compensation (typically in money) that a person would accept in lieu of receiving that change (willingness to accept). The benefits captured by this concept of value can be derived not only from good and services for which there are markets but also from non-market goods and services. In addition, both use and non-use (e.g., existence) values are included. Thus, economic valuation captures values that extend well-beyond commercial or market values. However, it does not capture non-anthropocentric values (e.g., biocentric values) and values based on the deontological concept of intrinsic rights. In addition, both willingness-to-pay and willingness-to-accept measures depend on the individual's current income (as well as market prices), implying that individuals with higher incomes will typically have higher benefits. This is viewed by many as a drawback of this approach to defining value.

There are multiple economic valuation methods that can be used in principle to estimate willingness to pay. These include methods based on observed behavior (market-based and revealed preference methods) and methods based on information elicited from surveys (stated preference methods). In contrast, in general measures of willingness to accept can only be obtained using stated preference methods.

Market-based methods seek to use information about market prices (or market demand) to infer values related to changes in marketed goods and services. For example, when ecological changes lead to a small change in timber or commercial fishing harvests, the market price of timber or fish can be used as a measure of willingness to pay for that change. If the change is large, then the current market price alone is not sufficient to determine value; rather, the demand for timber or fish at various prices must be used to determine willingness to pay for the change. In general, market-based methods are limited to valuing "provisioning" services supplied in well-functioning markets.

Revealed preference methods exploit the relationship between some forms of individual behavior (e.g., visiting a lake or buying a house) and associated environmental attributes (e.g., of the lake or the house). For example, travel cost methods (including applications using random utility models) use information about how much people implicitly or explicitly pay to visit locations with specific environmental attributes (e.g., specific levels of ecosystem services) to infer how much they value changes in those attributes. Hedonic methods use information about how much people pay for houses with specific environmental attributes (e.g., visibility, proximity to amenities or disamenities) to infer how much they value changes in those attributes. In contrast, averting behavior methods use observations on how much people spend to avoid adverse (environmental) effects to infer how much they value or are willing to pay for the improvements those expenditures yield.

In contrast to revealed preference methods, stated preference methods infer values or benefits from survey responses. In some cases, survey questions directly elicit information about willingness-to-pay (or accept), while under some survey designs (e.g., conjoint or contingent behavior designs) monetary measures of benefits are not revealed directly. Rather, some form of quantitative analysis is needed to derive benefit measures from responses to questions. Although the use of stated preference methods for

environmental valuation has been controversial, there is considerable evidence that the hypothetical responses in these surveys provide useful evidence regarding values.

While in principle the above methods can provide estimates of economic benefits, in some contexts use of them may be limited by time, resource, or data constraints. In such cases, proxies for or indicators of economic benefits can provide information about willingness to pay. Ecosystem Benefit Indicators (EBIs) offer a quantitative but not monetary way to illustrate ecological benefits in a specific setting. They use geo-spatial data to provide information related to the demand for, supply (or scarcity) of, and complements to particular ecosystem services across a given landscape based on social and biophysical features that influence (positively or negatively) the contributions of ecosystem services to human well-being. Examples of indicators include percentage of a watershed in a particular land use or of a particular land type, number of users of a service (e.g., water or recreation) within a given area, and distance to nearest vulnerable community.

4.4. Deliberative Processes that Involve Values

Most of the valuation methods described thus far involve a process under which information flows primarily in one direction (from the person whose values are being elicited to the analyst who is seeking to measure those values). In contrast, in some cases valuation is part of a deliberative process or facilitated interaction among decision makers, analysts, and stakeholder representatives or other interested parties that occurs over a more extended period of time (e.g., days or weeks). In such cases, the process itself becomes an important component for understanding and conveying information about values. These processes seek to elicit or incorporate information about all possible sources of value. In addition, they involve directly confronting tradeoffs that inevitably arise. Two examples of deliberative processes are decision-aiding processes and mediated modeling.

Decision aiding processes have been developed by decision scientists and applied in a number of contexts, including contexts involving environmental choices. From the perspective of decision science, valuation is not a separate exercise that then feeds into a decision made by others. Rather, it is part of a process designed both to discover values and, in many cases, ultimately to make policy decisions. This is based on the premise

that people's preferences and values for complex, unfamiliar goods (such as many ecosystem services) are often constructed during the process of elicitation and are multi-dimensional. This premise is in contrast to the premise underlying some of the methods discussed previously, most notably economic valuation methods, which assume that preferences are given and that values or benefits can be measured using a single metric such as willingness to pay.

Decision aiding processes can be applied either in a decision making context or an evaluative context. In either case, they involve a number of steps, including identification of the objectives of the process, definition of the attributes that will be used to judge progress toward the objectives, specification of the set of management options, measurement of changes in relevant attributes that would be realized under alternative management options, etc. (I'm not sure about these steps. KS) These steps draw on inputs from a variety of disciplines, such as economics, ecology, psychology, and sociology. The final output is either the selection or identification of a preferred management option (if the context is decision making) or a judgment about the current state of the system relative to a previous state (if the context is evaluative). Generally, the objectives that guide these final outputs are diverse and often multi-dimensional. Examples include maintaining some requisite level of ecological services, protecting endangered or threatened species, production of outputs such as resource extraction, tourism, and recreation opportunities, and supplying a sense of pride or awe (Gregory et al. 2001).

With mediated modeling the deliberative process focuses on stakeholder interactive development of a model representing a particular environmental system of interest. Stakeholder participation in model development occurs at all stages of the modeling process, from initial problem scoping to model development, implementation, and use. The models that are developed can be at any geographical scale, from watersheds or specific ecosystems to large regions or even the globe. The output of the process is a model that can be used to evaluate alternative scenarios or options of interest to those stakeholders. Most importantly, the model and the results derived from it have stakeholder buy-in and reflect group consensus. If the model is used to consider tradeoffs, then values must be explicitly incorporated into the model through the

specification of related parameter values. (Not sure about this. KS) These values can be drawn from other valuation exercises or based on other information that relates to value (e.g., use data).

4.5. Public and Group Expressions of Value

Traditional economic valuation methods attempt to measure and aggregate the values that individuals place on changes in ecological systems and services based on their personal preferences as consumers of those systems and services. An alternative approach is to try to measure the values that groups of individuals place on changes in such systems and services explicitly in their role as citizens – social/civic valuation. This approach measures the monetary value that groups place on changes in the systems and services when asked to evaluate how much *the public as a whole* should pay for increases in such systems and services (public willingness to pay) or should accept in compensation for reductions in the systems and services (public willingness to accept). The value measurement purposefully seeks to assess the full "public regardedness" value, if any, that the group attaches to any increase in community well-being attributable to changes in the relevant systems and services.

Social/civic values, like values based on personal preferences, can be measured either through revealed behavior or through stated valuations. One principal source of revealed values for changes in ecological systems and services are votes on public referendum and initiative involving environmental decisions. Other public decisions, however, also may provide measures of social/civil values, including official community decisions to accept compensation for permitting environmental damage, and jury awards in cases involving damage to natural resources. Where revealed values are difficult or impossible to obtain, social/civil values also can be measured by asking "citizen valuation juries" or other representative groups the value that they, as citizens, place on changes in particular ecological systems or services.

Analyses of the outcomes of referenda or initiatives (with or without a follow-up survey) seek to determine, for example, if the majority of the voting population feel that a given environmental improvement is "worth" what it will cost the relevant government body, given a particular means of financing the associated expenditure. Similarly, analyses of public votes about whether to accept an environmental degradation (e.g.,

through hosting a noxious facility) seek to determine if the majority of the voting population in that community feel that the environmental services that would be lost are "worth" less than the benefits the community would realize in the form of tax revenues, jobs, monetary compensation, etc. These approaches provide information about the policy preferences of the median voter, and under certain conditions can provide information about the mean valuations of those who participate in the voting process. The logic of using formal public outcomes to infer how much society values particular outcomes has been used previously to estimate the public's willingness to pay (in the form of a commitment of public expenditure) to reduce mortality rates from health and safety risks.

Like initiatives and referenda, citizen valuation juries provide information on social/civic values, but they measure stated rather than revealed value, and they incorporate elements of the "deliberative valuation" processes. The group is given extensive information and, after extensive discussion, is usually asked to agree on a common value or make a group decision. To date, citizen juries have typically been asked to develop a ranking of alternative options for achieving a given goal. However, a jury could also be asked to generate a value for how much the public would (or should) be willing to pay for a possible environmental improvement, or, conversely, how much it should be willing to accept for an environmental degradation. Experience with the use of citizen juries for ecological valuation is very limited to date.

4.6. Methods Using Cost as a Proxy for Value

A fundamental principle in economics is the distinction between benefits and costs. Benefits reflect what is gained by increasing the amount of a given (ecological) good or service, while costs reflect what must be given up for that increase. The value of ecological goods and services relates to the benefit side of this distinction. Nonetheless, some have used monetary measurements of production costs as measures of value. Examples include valuing habitat by the cost of obtaining or producing an equivalent habitat elsewhere (known as Habitat Equivalency Analysis (HEA)), valuing pollution reductions by the cost of a tradable emissions permit (which reflects the cost of producing one less unit of pollution abatement), and valuing a water source by the cost of producing or providing a replacement source.

It is well-known in economics that the cost of producing a good or its equivalent provides information about benefits and hence the value of that production only under very specific and limited conditions. In particular, a necessary condition for using cost or expenditure based estimates for purposes of benefit estimation is that the researcher *observes* people (or groups) revealing their preferences by *actually (and voluntarily) incurring costs* to avert (or tolerate) the environmental disruption in question. While this condition holds when, for example, individuals make voluntary expenditures on to reduce negative environmental health effects (e.g., from drinking contaminated water), it does not hold when expenditure decisions are pre-determined or constrained by other factors. In this latter case, there can be no presumption that benefits exceed costs, and hence costs provide no information about the associated benefits or value.

Applying this fundamental principle to ecological valuation implies that methods such as HEA yield, at best, the cost of cleanup or replacement, *not the value* of that cleanup or replacement, and certainly not the value of protecting the original ecology that was damaged (or might be damaged in the future). For this reason, the committee strongly rejects the use of cost-based estimates as a means of measuring benefits or value, except in the limited cases where the necessary and sufficient conditions for this to be an appropriate proxy for benefits hold.

5. ECONOMIC VALUATION FOR NATIONAL RULEMAKING

5.1. Introduction

The objective of this section is to examine the valuation of ecosystem services by the Agency in the context of the evaluation of the benefits and costs of national rules promulgated by the Agency and to make recommendations as to how ecosystem service valuation can be improved.

Most of the environmental laws administered by the Agency require that regulations be defined by a set of criteria other than benefits and costs. Indeed in some cases the legislation explicitly precludes consideration of costs or benefits in the standard setting process. For example in the case of the Clean Air Act, rules to establish primary ambient air quality standards for criteria air pollutants are meant to protect human health with an adequate margin of safety. Even in those cases where the law allows consideration of the benefits and costs, e.g., Safe Drinking Water Act, adherence to a strict "benefits must exceed costs" criterion is not required.

Nonetheless, an assessment of the benefits and costs of EPA actions play an important role in the context of national rule making for a number of reasons. First, analyses of Agency Rulemakings are required under the terms of Executive Order 12866 which states, "Each agency shall assess both the costs and the benefits of the intended regulations, and …, propose or adopt a regulation only upon a reasoned determination that the benefits of the intended regulation justify its costs" (Executive Order 12866, October 4, 1993). These assessments are commonly referred to as regulatory impact assessments or RIAs. They generally evaluate the form and stringency of the rules that are established to meet some other objective such as protection of human health. Second, an assessment of benefits and costs can be mandated by law. For example, the prospective analysis of the benefits of the Clean Air Act Amendments of 1990 was mandated by Section 812 of the Clean Air Act Amendments of 1990, which requires the Agency to develop periodic Reports to Congress that estimate the benefits and costs of various provisions of the Clean Air Act. Finally, the benefit-cost analyses developed in national rulemaking may later be taken into account by executive-branch officials and legislators in formulating and proposing new national rules or for other purposes.

Therefore, a complete, accurate, and credible benefit-cost analysis of a given rule can have broad impacts even if the analysis does not determine whether the current rule is enacted.

Circular A-4 from the Office of Management and Budget (OMB, 2003) makes it clear that what is intended by Executive Order 12866 is an economic benefit-cost analysis for the proposed rules conducted in accordance with the methods and procedures of standard welfare economics. Thus, in the context of national rulemaking, the term "benefit-cost analysis" has a specific meaning. It means that to the extent possible the benefits associated with changes in goods and services or prices due to the rule are to be measured by the sum of the individuals' willingness to pay for them. Similarly, the costs associated with regulatory action are to be evaluated as the losses experienced by people and measured as the sum of their willingness to accept compensation for those losses. Thus, a benefit-cost analysis begins with a specification of what environmental conditions would be throughout the areas affected by the rule with and without it. These changes are then evaluated based on individual willingness to pay and to accept compensation and aggregated over the people (or households) experiencing them.Circular A-4 includes recognition that it might not be possible to express all benefits and costs in monetary terms. In these cases, it calls for measurement of these effects in physical terms. If that is not possible, there should still be a qualitative description of the benefits and costs (OMB, 2003, p. 10).

This section considers ecological valuation in the context of national rulemaking governed by Executive Order 12866 and OMB's Circular A-4. It thus focuses on the use of economic valuation methods that seek to monetize benefits based on the concept of willingness to pay (or accept compensation), recognizing that when monetization is not possible, the Agency should seek to quantify impacts in biophysical terms or provide a science-based qualitative description as prescribed by Circular A-4. As background for this discussion, the committee examined three specific examples of previous agency benefits assessments: (1) the agency's benefit assessment for the final effluent guidelines for the aquaculture or the concentrated aquatic animal production industry (US EPA 2004), (2) its assessment for the recent rulemaking regarding concentrated animal feeding operations (CAFOs) (US EPA 2002), and (3) the prospective analysis of the benefits of

the Clean Air Act Amendments of 1990 (US EPA 1999). [30] Brief descriptions of the three benefit analyses are presented in separate text boxes. These examples provided insights that are reflected in the discussion and recommendations throughout this section.

5.2. Implementing the Proposed Approach

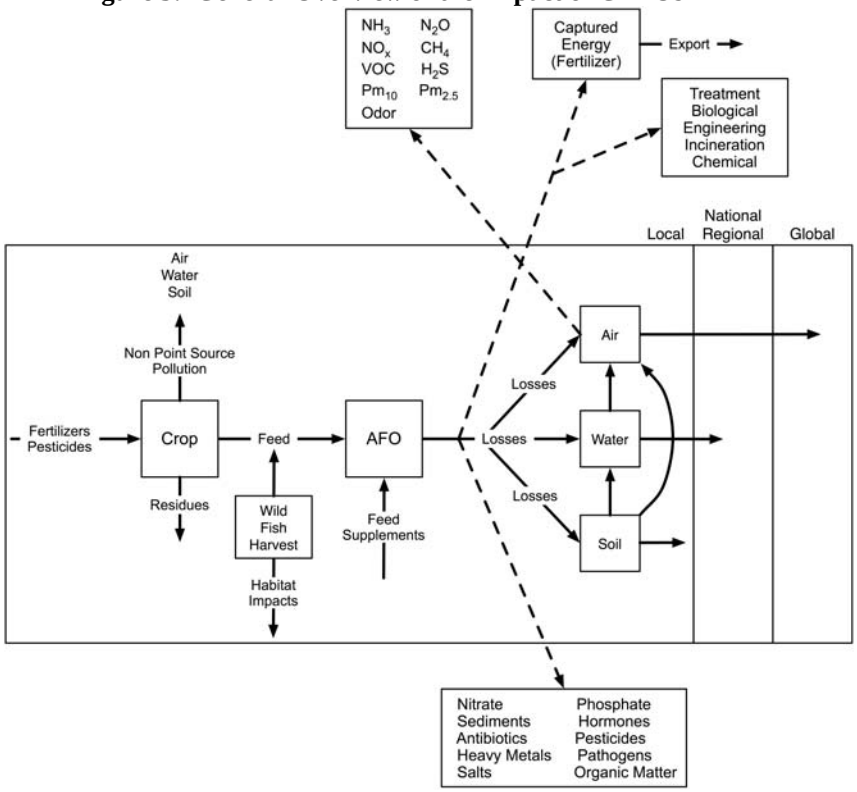
This section describes how EPA could implement the integrated and expanded approach to ecological valuation proposed in this report in the context of national rulemaking and benefit-cost analysis. It focuses on three components of that process (see Figure 2): a) identification of socially important assessment endpoints; b) prediction/quantification of bio-physical changes in assessment endpoints; and c) characterization of the value of changes in monetary and non-monetary terms. Each is discussed in turn.

5.2.1 Identification of socially important assessment endpoints

Identification of socially important assessment endpoints requires information about both the potential biophysical effects of the Agency's action and the ecological services that matter to people. To guide the collection of this information, the Agency should develop a conceptual model of the ecological and economic system being analyzed. Conceptual models can allow the Agency to take a broad view of the complexities involved in addressing ecological changes (see discussion in section 3? above). Determination of the important ecological effects could draw on technical studies of impacts and their magnitudes, as well as solicitation of expert opinion regarding the nature of physical and biological effects of a regulatory change. As an example, Figure 5 gives a general overview of the ecological impacts of CAFOs, which enables a comprehensive evaluation of what is happening to the environment and where the levers are for improving environmental performance. (Hal? What is the source for this figure? KS) This overview could be used to develop a conceptual model that identifies potential ecological services that might be affected by CAFO regulation. It should be standard practice for the Agency to develop such a conceptual model before other analytical work begins on a benefit assessment or RIA. The analytical blueprint required as part of EPA's process for developing rules should call for development of a conceptual model

for ecological valuation and specify the interdisciplinary team to be involved in developing it.

Figure 5: General Overview of the Impact of CAFOs



As noted, the conceptual model should include information about the changes that are likely to be of greatest concern to people. The committee believes that identification of what matters to people cannot be done deductively. Rather, it requires an examination of the evidence gleaned from a variety of research approaches. It is important to distinguish the processes used to enumerate the goods and services that are important to people from the process used to evaluate benefits and costs. Where the analysis is being conducted to meet a mandate for benefit-cost analysis (as is the case for RIAs), the computation of benefits and costs must be consistent with the methodological requirements of the benefit-cost framework. However, the process of identifying early on the public concerns associated with a given rule can be undertaken with a variety of methods.

The suite of methods that can be used to assess public concerns includes surveys, public meetings, focus groups, content analysis of public comments, and so forth.

Relevant initiatives, referenda, or community decisions might also be available in some jurisdictions to get a more robust indication of the preference for cleaner water and/or the avoidance of the various risks. More specifically, possible approaches for obtaining information about public concerns include:

- Inventory of the reasons invoked in similar rulemaking processes in other jurisdictions (e.g., state and local).
- Inventory of the concerns expressed in public hearings (perhaps with weightings based on the frequency of concerns raised). For example, local vs. national concerns can be quantified through content analysis of transcripts. Where local debates over allowing fish farming have occurred, the discourse could reveal what people care about.
- Focus groups and surveys of concerns (can be lists of concerns, or quantified by ranking priorities).

In addition, the committee suggests that EPA experiment with holding open meetings for the public and agency staff to develop a conceptual model for a particular rulemaking and consider using mediated modeling methods for these meetings. Such an approach would provide an interactive forum for determining the ecological changes that are important both biophysically and socially.

Whatever methods are chosen, to increase transparency related to the analytical process for developing rulemaking documents, the Agency should document in its benefit assessments and RIAs how "socially important assessment endpoints" were identified for the analysis. It should clearly identify the criteria for including effects within the core analysis and how these criteria were applied to those analytical choices. In addition, EPA should specifically document in final benefit assessments and RIAs how the Agency incorporated relevant input on ecological values related to the rule from public meetings on the proposed rule. It would also be helpful to provide a specific section in RIAs and benefit assessments describing how the Agency addressed comments related to ecological values and valuation if specific comments related to those issues were submitted as part of the formal notice and comment process for draft benefit assessments and RIAs.

Finally, the analytical blueprints and conceptual model that was used to guide the analysis should be part of the public record for every rulemaking and available on-line.

5.2.2 Bio-physical prediction of changes in assessment endpoints

Once the relevant assessment endpoints have been identified, Circular A-4 requires that the changes in those endpoints be quantified (to the extent possible) either as an input into monetization or as an indicator of importance when monetization is not possible. Conceptually, this requires first a prediction in the changes in stressor resulting from the rule and then a prediction of how the change in stressors will likely change the assessment endpoints or services.

In many rulemaking contexts, it is difficult to predict even the changes in stressors, let alone the resulting impact on endpoints. For example, in the RIA for the aquaculture rule, it was difficult to quantify the changes in stressors because baseline data on stressor levels were not available and the rule did not impose quantitative maximum discharge levels. (So what did they do? Can we give any advice on this? KS) In addition, in the past the agency has generally chosen to focus on stressors whose impacts can be monetized with readily available techniques and/or estimates from the existing literature. All three of the rulemaking benefit assessments that the committee reviewed provide evidence of this. For example, for the aquaculture rule, the agency used the QUAL2E model to predict ecological impacts. While this model can estimate the interactions among nutrients, algal growth and dissolved oxygen, it is not capable of ascertaining the impacts of total suspended solids, metals, organics, etc. on the benthos and the resulting cascading effects on aquatic communities. The choice of QUAL2E appears to have been driven largely by the ability to link its outputs with existing estimates of willingness to pay for water quality improvements taken from Carson and Mitchell. Rather than choosing stressors based on the ability to readily monetize their impacts, the agency should use the conceptual model (see discussion above) to guide the selection of stressors, and then seek to use a suite of ecological models that can predict the impacts of changes in these stressors on a broader set of the relevant assessment endpoints.

Quantifying changes in assessment endpoints is particularly challenging in national rulemaking contexts, and there are many issues that need to be addressed in

order to establish a convincing analysis of the benefits of a national rule. . Assessment at the national level means that impacts can be both varied and potentially large, implying the need for a more comprehensive analysis. Yet comprehensive analysis is particularly difficult precisely because of this scale and the associated complexity. For example, the committee's review of the CAFO rulemaking noted the following issues that stem from the varied and complex environmental consequences of CAFOs (see Figure 5):

- Multi-media effects, i.e., interrelated impacts on both water and air quality;31
- Impacts across multiple geographical scales (e.g., local, regional, global);32
- Differences in the time persistence of pollutants (e.g., days vs. decades);33
- Clustering and the need for site-specific analysis due to uniqueness of site characteristics associated with impacts;34 and
- Ecological impacts through supply-chain effects that are geographically dispersed.35

Some of the links between stressors and endpoints are well-understood and relatively easily quantified. Examples include the movement of phosphorus and nitrogen from manure into surrounding waters. Phosphorus in particular has been studied intensively and, importantly, its impact has been well demonstrated by whole ecosystem experiments for fresh water.[36] Similarly, species that the public or experts particularly value have been studied in sufficient detail that there are process models of production and interaction with other species. Examples include a corn plant and a salmon. Scientists can specify a production function for these organisms and use that function to predict the impact of changes in stressors.

However, many of the links between stressors and assessment endpoints are: a) not fully understood scientifically, and/or b) not fully appreciated by the public. For example, one of the important ecosystem services affected by the CAFO rule is the support of populations of fish species that are targets of recreational angling. To predict the effects of the rule on ecosystem services, one would need to know how populations of

these species change and how population changes affect anglers' success rates. These links are not well understood at the level required for a comprehensive national analysis. Scientific knowledge is especially lacking in understanding the ecological impacts of substances such as heavy metals, hormones, antibiotics, and pesticides. Yet these substances can have important and far-ranging impacts that could be significant at the national level. For example, arsenic in poultry manure moves into local environments as well as through different pathways to places more distant, either through the sales of incinerator ash for fertilizers from fuel generators (I don't follow this. Do you mean they burn the manure for fuel and then sell the ash for fertilizer? KS), or directly by the use of dried and pelletized manure in places distant from the source (Nachman, et al., 2005).

There are many things that are well known scientifically, yet the general public is not fully aware of and hence has no appreciation of or informed opinion about them. For example, the full chain of connections in the production of animals in CAFOs is indicated in Figure 5 is not generally understood or appreciated by the public. Similarly, the public does not generally understand the organisms and processes involved in breaking down waste products and the resulting services provided. For example, certain groups of soil organisms maintain soil structure by their burrowing activities, while other kinds of organisms shred the organic material into smaller units that are in turn utilized by microbes that release nutrients in a form that can be utilized by higher plants for their growth. However, the general public has little appreciation for the "services" these organisms provide (e.g., Weslawski, et al. 2004). Again, this might be exacerbated in national level analyses where ecological impacts and vulnerabilities can vary substantially across locations.

The combination of variation, complexity, and gaps in information and understanding make it difficult for the Agency to assess the ecological impacts of its actions, particularly at the national scale. As noted above, Circular A-4 requires the agency to monetize impacts that can be monetized, quantify those that cannot be monetized but can be quantified, and describe qualitatively (based on scientifically-credible theories or evidence) impacts that cannot be quantified. The actual process for implementing the Circular, however, requires a reversal of this order, namely, first

impacts should be described or characterized qualitatively, followed by quantification and ultimately monetization where possible.

As noted above, characterization of ecological impacts requires a conceptual model (see detailed discussion in Section 3). Such a model would link the various levels of organizations of ecosystems that are involved in the provision of ecosystem services, as illustrated in Figures 3 and 4 in Section 2. This model can be used as the basis for a qualitative but detailed description of the ecological impacts of a given change. Table 4, taken from a recent report on climate change by the Union of Concerned Scientists and the Ecological Society of American, provides an example of how ecological impacts can be summarized in a detailed, qualitative form. (I think this example is useful, but there may be better examples. Can we construct a table like this for the CAFO or aquaculture rule? Would it be any different from Exhibit ES-4 in the CAFO report? KS) However, just a listing that summarizes possible impacts is not sufficient. Such a summary should be accompanied by justification based on the conceptual model and the associated theoretical and empirical scientific literature. To the extent possible, the existing literature should be used to draw inferences about the likely magnitude or importance of different effects, even if only qualitatively (e.g., high, medium, low).

Table 4: Impacts of Climate Change on Stream Ecosystems (Source: Kling, et al., 2003, "Confronting Climate Change in the Great Lakes Region: Impacts on our Communities and Ecosystems," Union of Concerned Scientists, Cambridge, MA, and Ecological Society of America, Washington, D.C., p. 26

Climate-Driven Change	**Likely Impacts on Physical and Chemical Properties**	**Likely Impacts on Ecosystem Properties**	**Intensifying or Confounding Factors**
Earlier ice-out And snow melt	Peak flows occur earlier. Ephemeral streams dry earlier in the season Backwater pools experience anoxia earlier.	The timing of fish and insect life cycles could be disrupted.	Snowmelt occurs earlier and faster in urban areas and where coniferous forest harvest has occurred.
Lower summer water levels	More headwater streams dry; more perennial streams become intermittent. Concentrations of dissolved organic carbon decrease, thereby reducing ultraviolet-B attenuation. Groundwater recharge is reduced.	Habitat decreases in extent. Hydrologic connections to the riparian zone are reduced. Groundwater recharge is reduced. Species with resting life stages or rapid colonizers dominate communities	Impervious surfaces and impervious soils exacerbate stream drying due to reduction in infiltration and groundwater recharge.
More precipi-tation in winter and spring and increased water levels	Spring floods reach greater heights. Surface runoff increases. Nutrient and sediment retention decrease. Groundwater recharge potential increases.	Floodplain habitat for fish and invertebrates grows. Hydrologic connections with wetlands increase.	Precipitation occurring when soils are frozen results in higher runoff and increases flood height.
Etc…Add rest of table…..			

To move from a qualitative to a quantitative prediction of impacts, the conceptual model must be linked with one or more ecological models that capture the essential linkages embodied in the conceptual model and are parameterized to reflect the range of relevant scales and regions. Criteria for choosing among alternative models were discussed in Section 3. The objective is to use the models to generate metrics to compare biological conditions with and without the rule to see the potential effect of the rule on the delivery of ecosystem services.

Some ecological models predict certain processes of entire ecosystems, such as primary productivity and ecosystem water balance. For these types of models simple features of the ecosystem are sufficient (for what?? KS), such as height of canopy of the vegetation and leaf and rooting type in addition to local physical drivers of these processes. (Can we say more about the use of these types of models, particularly for large scale, i.e., national, analyses? Examples of how they could be used? KS)

However, as discussed in Section 2, because of the complexity of ecosystems, ecosystem models often focus on functional groups that combine to deliver ecosystem services. These groupings reflect the fact that species share similar functions and can be grouped accordingly. Ecosystems can then be categorized not by species but by the abundance of the various functional groups present, e.g. decomposers of various kinds, as exemplified by the array of nematode types mentioned earlier, or bacteria that are specific to producing a particular breakdown product, such as the methane producers and those that produce nitrous oxides. General models can be used to predict the responses of these functional groups to environmental change, which can often indicate impairment of certain ecosystem processes and hence changes in ecosystem services. In general, this requires consideration of multiple functional levels. For example, in the context of the CAFO rule, there are actually many ecosystem types, which are spatially separated, that need to be considered. The soil system is the initial target of stressor impact where the main focus of analysis should be on the various soil macro-organisms involved the initial breakdown of organic matter as well as on those microbial types that finish the job of converting the organic matter to various end-products. As these products, or in some cases organic matter directly, flow into free-flowing water systems, then different kinds of organisms become important in selecting indicators of ecosystem function impairment.

There are readily available and fully tested techniques for evaluating different functional groups and in theory metrics related to these groups could be used to quantify the ecological impacts of a given rule. Specifically, the abundance of these groupings can be readily quantified in any before-and- after rule condition. For example, at the base of the ecosystem is its potential and realized biological diversity. Thus metrics that look at the impact of the rule on species richness and various diversity indices get at this directly. (Hal, can we give some other examples? Are we talking about predicting the number of nematodes or bacteria with and without the rule, or is there something else you have in mind here? How would this be applied at the national level? KS)

The key though is to identify those components of each of the functional levels that are most directly related to the services of interest and thus provide ecological indicators of the state of the system in relation to the change in stress level. There are a number of approaches to limiting the indicators to those that will provide the most direct information relevant to the services in question. One is to focus on those functional groups that play a most prominent role in service provision. (Can we give some examples of indicators that are closely linked to services? For example, UCS/ESA report climate change impacts in terms of changes in "total annual stream flow into reservoirs", "number of years with dry or critically low stream flow conditions," etc. Are there some metrics or indicators like this that we might want to give as examples that could be applied at the national level?

Also, what about Jim Boyd's Ecological Benefit Indicators? Can we use examples from there as a way to "quantify when you can't monetize"? Is this feasible at the national level? KS)

In summary, the initial conceptual model of a system provides the big picture of the possible environmental impacts of the rule. Then, when focusing on just the outputs from specific facilities such as CAFOs or aquaculture facilities that are covered in a rule, there is a large array of potential metrics that would indicate the success of rulemaking in providing better ecosystem services to society. (Hal: Can we provide some examples of metrics that can used in national analyses? KS) In addition to looking at end point services only, it is important to look at the ecosystem service providers, even though they cannot be directly monetized. The suggestion here is through an analysis of the

structures of the systems that are impacted it should be possible to focus on functional types that are most directly involved in providing the services in question. There are ample tools available for making these measurements. (Hal: Again, some examples of these tools showing applications that could be used in national analyses would be really helpful. KS)

5.2.3 Monetary Measures of Value

Circular A-4 calls for the monetization of benefits whenever possible. Although there are a variety of methods that can be used to determine values for purposes of identifying socially relevant assessment endpoints (see discussion above) and for value assessments in other contexts (see Sections 6 and 7), in the context of benefit-cost analysis the only approach to monetization consistent with the premises underlying this analysis is the use of economic valuation methods. The inclusion of numerical values based on other methods such as those mentioned above, even if measured in dollar terms, is problematic because it implies adding together numbers that are based on quite different methods, assumptions, and underlying premises. Thus, for both theoretical and empirical consistency, the measure of benefits in a benefit-cost analysis should be based on economic valuation.

There is a large and growing theoretical and empirical literature within economics on methods for assigning monetary values to environmental changes. These methods use either observed behavior (revealed preference) or responses to surveys (stated preference or contingent valuation/choice) to estimate willingness to pay (or accept compensation) for these changes. While there have been controversies surrounding the use of these methods, particularly the stated preference methods, existing research supports the view that, when appropriately used, these methods can provide reliable estimates of willingness to pay (see related discussions in Part 3, section ? and Appendix A). (Since OMB seems skeptical about the use of CV, I think we want to say that the committee thinks it is appropriate to use (good) CV estimates in benefit assessments. I think this is, in fact, what prompted us to ask Jon and Allyson to prepare Appendix A. KS)

Nonetheless, the Agency needs to ensure that the call for monetization does not unduly restrict the types of ecosystem impacts considered in the benefit assessment, or

lead to inappropriate application of economic valuation methods (including benefits transfer). As noted above, the call for monetization has often driven Agency decisions regarding the focus of ecological benefit assessments. This applies not only to the types of ecosystem services included in the detailed assessments but also the ecological models used to predict biophysical impacts. For example, the Agency's assessment of the aquaculture rule focused primarily (almost exclusively??) on recreational impacts and used the QUAL2E water quality model to predict the changes in several water quality indices that would result from implementation of the rule. The choice of QUAL2E appears to have been driven largely by the ability to link its outputs with readily available, off-the-shelf monetary estimates of willingness-to-pay for changes in water quality indices taken from the Carson-Mitchell contingent valuation (CV) study. The principal advantage of this approach is that it utilizes a study designed to be national in scope and has a simple willingness-to-pay relationship that allows the analysis to be done relatively quickly, without new research and the associated significant expenditures on research resources. Also, it can be applied using a straightforward conceptual logic that is easy to understand. However, use of the Carson-Mitchell estimates has a number of limitations that raise concerns about the resulting benefit estimates. Most notably, the study was conducted more than 20 years ago, it was designed for a different purpose and was not intended to apply to specific rivers or lakes, the water quality index used was highly simplified, and the index it used was never designed to reflect ecological services related to water quality (other than those related to fish). Thus, in an effort to focus on effects that could be readily monetized, the Agency appears to have limited both the types of services considered and the ecological and economic models used to estimate the impacts of the rule on those services.

The previous section discusses the need to consider a broad range of ecosystem services when assessing the benefits of national rules, even if the benefits associated with changes in those services cannot be monetized. However, even when the benefits can be monetized, the above example highlights the need for appropriate application of monetization techniques to ensure scientifically-credible benefit estimates. In many cases, time and resource constraints will necessitate use of benefits transfer. However,

care must be taken to ensure that the benefit estimates that are used are appropriate (see discussion of benefits transfer in Part 3, section ?).

There have been individual studies of recreational angling in specific areas relating the choice of recreation site to measures of travel cost and proxy measures for the availability of specific fish populations that could serve as a basis for benefit transfer. Most of this work has focused at marine fishing. However, some studies have been undertaken for freshwater systems. [37] For example, when estimating the recreational benefits of reducing acid deposition in Adirondacks lakes, the Agency used a fairly recent published study of recreational angling choices of households in New York, New Hampshire, Maine, and Vermont (Montgomery and Needelman, 1997). (Do we think this was a good application of benefit transfer? KS) This was a random utility model of site choice. Measured pH of lakes was used as an indicator of the level of ecological services from each lake. The literature on the economics of recreational angling shows that likelihood of success as measured by numbers of fish caught is a major determinant of demand for recreational angling (see Phaneuf and Smith [2005] and Freeman [1995] for reviews). To the extent that populations of target species are correlated with pH levels, pH will be a satisfactory proxy for fish populations and angling success rates.

There are several types of alternative models that can also be used that would allow direct use of the outputs of some type of ecological model for ecological impacts. One possibility would be to use an existing model linking the physical descriptors of water quality to recreation behavior to estimate the benefits per trip for a change in water quality conditions comparable to the rule's effect, had it been experienced in each of the areas. (I'm not sure I have this right. I was confused by the original sentence. KS) These estimates could then be used in a summary or meta function describing how the local choice set of recreation sites and economic characteristics of the recreationists as well as the character of the changes from existing baseline conditions influenced the estimates of unit benefits. Such a meta function could then be considered for other areas. (references??) Alternatively, the models could be adapted to be directly applied to choice sets composed for affected areas. In this case the recreation behavior necessary to operationalize the model could be extracted for some of the areas from EPA's National Survey on Recreation and the Environment (NSRE) for 2000 and 2004. The logic

involved has two key steps: a) translation of the effect of the rule for a set of local water quality conditions that is matched to some set of economic behavior for that area that is influenced by the water quality; and b) adaptation of an economic model of tradeoffs people would be willing to make to improve one or more aspects of the water quality for the area so that economic and ecological factors affecting the tradeoffs are represented in the summary function. There is precedent in the literature on benefits transfer for these types of analyses (see Rosenberger and Loomis [2003] and Navrud [in press], for examples of how this logic might be used in benefits transfer).

A second class of models for evaluating stressors affected by the rule are the stated preference and stated choice models that highlight water quality attributes. While the record here is not as extensive as it is for the revealed preference random utility (RUM) models, there are several candidate studies (references??). These analyses are based on surveys that elicit respondent choices among a set of options, plans for reducing effluents or for improving water quality defined in terms of pollutants and or characteristics of ecosystems. The logic is comparable to that described for the RUM. The effects of the rule need to be adapted to the features of each of the models and projected unit benefits derived. Then the factors affecting the benefit measure for each are used with a model in a summary analysis that can facilitate transfer to areas that do not have such models but are affected by the rule.

In addition to recreational impacts, some ecological services affect the well-being of homeowners living near the ecological systems providing these services. Examples include water regulation and flood control and the amenities associated with healthy populations of plants and animals. Residents' willingnesses to pay for these services can be capitalized into housing prices. The hedonic property value method can be used to obtain estimates of the values of these services. For examples, see Leggett and Bockstael (2000), Mahan, et al. (2000), Netusil (2005), and Poulos, et al. (2002). These estimates could then be candidates for use in a benefits transfer.

Can we add something somewhere about doing top down vs. bottom up analyses? The following text refers to using a bottom up approach for the aquaculture rule. KS

A preferable approach for estimating values based on recreation activities would be to do site-specific revealed preference (travel cost or random utility model) or stated preference analyses for a set of representative sites and to aggregate the results of these models to the 242 sites affected by the rule. The difficulty in undertaking such an analysis stems from the limited regional character of the available applications. Often the affected areas represent very idiosyncratic local conditions and are not nationally generalizable.

Can we say something somewhere about the difficulty of doing original surveys, given time and resource constraints and requirements for OMB approval? KS

Can we add something somewhere about the use of databanks that can be used for benefit transfer or the investment in a "generic" survey comparable to Carson and Mitchell that could generate WTP estimates that might be able to be used for a variety of valuation applications? KS

Benefit assessments and RIA should feature prominent discussions of ecological services that describe how ecological services were identified and analytical choices were made to assess and report on changes in service flows. In addition, they should clearly identify the values that were a) monetized using economic valuation methods, b) quantified (but not monetized), and c) described qualitatively. However, rather than simply designating them as "non-monetized", as for example in the CAFO benefit assessment, we recommend that the non-monetized but quantified impacts be reported explicitly (in conjunction with the monetized benefits) measured in the units that make sense from a biological perspective, and that the non-quantifiable impacts be described in as much detail as is feasible. Furthermore, any summary listing of the benefits and costs should include all three types of benefits, with the monetized and quantified benefits measured in the appropriate units (dollars or biophysical units). When monetized benefits are aggregated, the resulting sum should always be described as the "Total Monetized Benefits" rather than the "Total Benefits". In the past, EPA has sometimes reflected the non-monetized benefits in aggregate measures of benefits by including an

entry in the summary table of benefits (and costs) such as +X or +B to indicate the unknown monetary value that should be added to benefits if the value could be determined. While such an approach indicates that the monetary measure of benefits (and costs, too, if appropriate) is not a complete measure, the +X or +B provides little information about the extent or nature of the under-estimation and can be easily over-looked when the results of the benefit assessment are used. Always designating the sum as "Total Monetized Benefits" provides a continual reminder of what is (or is not) included in this measure. In addition, always reporting total monetized benefits together with key quantified but non-monetized impacts measured in biophysical units provides a more accurate and complete indication of total benefits than a simply designating total benefits by the sum of the monetary estimates plus an unknown factor X or B.

5.2.4 Uncertainty Analysis

Because of the difficulties in both estimating biophysical impacts of an EPA rule and the associated benefits or costs, it is important that EPA characterize the uncertainty associated with its benefit assessment. For example, EPA should include a separate chapter on "Uncertainty Characterization" in each benefit assessment and RIA. The chapter should discuss the scope of the benefit assessment, the different sources of uncertainty [e.g., Biophysical Changes and their Impacts; social information about endpoints, valuation methods (including use of "benefit transfer")], and report on methods used to evaluate uncertainty. Within the section on "scope" the Agency should discuss the types of "socially important" values related to the issue that were included in the assessment and those that were excluded because they were not conceptually appropriate for the benefit assessment or RIA. At a minimum, the chapter should report ranges of value and statistical information about the nature of uncertainty for which data exist. For each type of uncertainty, information similar to that reported in the Agency's prospective analysis of the benefits and costs of the Clean Air Act Amendments (US EPA, 1999) should be reported and a summary of this information should appear in the executive summary of the RIA or Benefit Assessment. Specifically, EPA should report: a) potential source of error; b) the direction of potential bias for overall monetary benefits estimate; and c) the likely significance relative to key uncertainties in the overall monetary benefit estimate. More generally, benefit assessments and RIAs should

highlight in quantitative and qualitative terms any "socially important assessment endpoints" identified as appropriate for the analysis that were not monetized.

To stimulate the exploration and development of methods needed to enhance EPA's capacity for ecological valuation, EPA should seek, for each rulemaking, to conduct a satellite sensitivity analysis using different methods from the core analysis, and preferably appropriate innovative methods, for one or more components of the core analysis. Such a satellite sensitivity analysis would serve to develop experience with innovative methods and to test the results of findings in the core analysis . The plan for the sensitivity analysis should be discussed in the analytical blueprint for the benefit assessment or RIA or the rationale for not including the sensitivity analysis should be discussed in this document, which would be part of the public record for the rulemaking and available on line.

In addition, the committee advices the Agency to develop an extramural grant activity within the Science to Achieve Results (STAR) program focused on method development specifically for recurring rulemakings (e.g., for rulemaking associated with programs like EPA's National Ambient Air Quality Standards or Effluent Guideline programs). Such a focused effort could help develop methods for expanded applications of monetary and non-monetary methods for valuing ecological effects that will have foreseeable benefits for Agency regulatory programs addressing ecological protection issues. The Committee also advises the Agency to host annual Agency-wide meetings to discuss methods used in regulatory impact analyses and benefits assessments and methods needed for full characterization of the effects addressed by the regulatory actions associated with those efforts.

5.3. Conclusions and recommendations

a) A significant barrier to any kind of valuation is the lack of information on how the levels of ecosystem services would be affected by the rule. Reasons for this include:

- In some cases (e.g., requirements for best management practices, absence of baseline data), the changes in the levels of ecological stressors were not known.

- The models do not predict changes in the relevant ecosystem services. For example, the links between outputs of some ecological models and human uses of the ecosystem were not known (e.g., the relationship between changes in fish populations and changes in recreational angling).
- The lack of site specific ecological data.

b) Methods exist for estimating economic values for at least some ecosystem services. But applying these methods in the context of new regulations could require original research that is costly and time consuming.

c) The Executive Order that mandates a benefit-cost analysis for major rules adopts a national perspective. Thus analysts undertaking the research needed to prepare benefit-cost analyses have tended to favor models and or estimates that also have a national perspective. This so-called "top down approach has caused them to overlook the possibility of adapting a set of regional studies more closely aligned to the changes in the ecological effects so that these studies could meet the goals of a national analysis. This alternative "bottom-up" approach would proceed by establishing separate estimates for each regional grouping or group of similar facilities and then adding them together to obtain the national estimate.

d) Since economic values are context dependent, benefits transfer very likely requires a much larger set of value estimates than is currently available.

e) Methods exist for estimating non-economic values for at least some ecosystem services. But these methods do not have a role in a benefit cost analysis.

f) It is important to involve both ecologists and economists at the earliest stages in the development of an analytical plan for ecological benefits assessment.

g) There needs to be better communication between the program offices and the Agency's Office of Research and Development (ORD) concerning the research needs of the program offices and the resources available from ORD.

Text Box 1: The Aquaculture Effluent Guidelines

Title III of the Clean Water Act (CWA) gives EPA authority to issue effluent guidelines, national standards for wastewater discharges to surface waters and publicly owned treatment works (municipal sewage treatment plants). The standards are

technology-based (i.e. they are based on the performance of available treatment and control technologies). The proposed effluent guidelines that were analyzed would require that all applicable facilities prevent discharge of drugs and pesticides that have been spilled and minimize discharges of excess feed and develop a set of systems and procedures to minimize or eliminate discharges of various potential environmental stressors. The rule also includes additional qualitative requirements for flow through and recirculating discharge facilities and for open water system facilities.

For most of these requirements, it is not possible to specify the change in the levels of environmental stressors since the rule did not impose quantitative maximum discharge levels. In addition, for most of these stressors, baseline data on discharges in the absence of the rule were not available.

The Agency identified the following potential ecological stressors: solids; nutrients; biochemical oxygen demand from uneaten food and feces; metals (from feed additives, sanitation products, and machinery and equipment); food additives for coloration; feed contaminants (mostly organochlorides); drugs; pesticides; pathogens; and introduction of non-native species. Some of these (for example, drugs, pathogens) were thought by the Agency to be very small in magnitude and not requiring further analysis. To this list C-VPESS added habitat alteration from changes in water flows.

The Agency analyzed the effects of changes in these stressors on dissolved oxygen, biochemical oxygen demand, total suspended solids, and nutrients (nitrogen and phosphorus). There appear to have been two reasons why the remaining endpoints were not quantified:

- The lack of data on baseline stressor levels and how regulation would change these levels.
- The Agency didn't use a model capable of characterizing a wide range of ecological effects. The Agency used the QUAL2E rather than the available AQUATOX model. The choice of QUAL2E appears to have been driven largely by the ability to link its outputs with the Carson and Mitchell valuation model.

The Agency estimated benefits for recreational use of the waters and non-use values. To estimate these values, the Agency estimated changes in six water quality parameters for 30 mile stretches downstream from a set of representative facilities and calculated changes in a water quality index for each facility. The Agency then used an estimated willingness to pay function for changes in this index taken from Carson and Mitchell (1993). Carson and Mitchell had asked a national sample of respondents to state their willingness to pay for changes in a water quality index that would move the majority of water bodies in the United States from one level on a water quality ladder to another resulting in improvements that would make possible boating, fishing and swimming in successive steps. This contingent valuation survey was conducted in 1982-83 and was not intended to apply to specific rivers or lakes.

The aggregate willingness to pay for the change in the water quality index for each representative facility was then used to extrapolate to the population of facilities of each type affected by the rule.

Text Box 2: The CAFO Effluent Guidelines

Context:

In recent years there has been substantial growth of the livestock industry in the United States as well as in many other parts of the world. This growth has been characterized by a dramatic reduction in the number of farm operations producing livestock and a big increase in the number of animals per farm unit. Finally, there has been a geographic concentration of these intensive units, particularly in the Southeast and Mid-Atlantic states. Manure production in these intensive facilities simply exceeds the capacity of nearby farmland to utilize it in plant production, resulting in a major disposal issue and hence threat to ground and surface waters as well as a problem with local air pollution.

These structural changes in the industry led to the present CAFO rule that was issued in December of 2002. This rule focused on the largest operations that represent the greatest environmental threats. These units are required to implement comprehensive nutrient management plans and to submit annual reports summarizing their operations.

What are the environmental issues?

The manure from livestock operations produces a variety of potential pollutants which can migrate to ground water, streams, rivers, and lakes. These pollutants include nitrogen, phosphorus, sediments and organic matter, heavy metals, salts, hormones, antibiotics, pesticides and pathogens (over 150 pathogens that are found in manure are human health risks). Further CAFO facilities release a variety of gases and material into the atmosphere including particulates, methane, ammonia, hydrogen sulfide, odor causing compounds, and nitrogen oxides.

Of the water-polluting materials, which are covered in the CAFO rule, excess nutrients can cause direct impacts on human water supply through excess nitrates, impacts on agriculture through excess salts in irrigation waters, as well as eutrophication of water bodies, anoxia and toxic algal blooms. These latter effects can result in fundamental changes in the structure and functioning of aquatic ecosystems including cascading effects that reduce water quality and species diversity. Uncontrolled releases of animal wastes have resulted in massive fish mortality.

Pathogens in polluted waters are a health hazard both directly as well as through the food chain, for example crops and shellfish. The potential impacts human health impacts of antibiotics and hormones in wastes have not been well identified but are of concern.

<u>How were the environmental impacts quantified?</u>

Of all of the potential environmental impacts, the CAFO benefits analysis focused to a large extent on the nutrient runoff from land where manure has been applied quantifying the benefits that would accrue from the manure management requirements of the CAFO rule. To do so they utilized the GLEAMS model (Groundwater Loading Effects of Agricultural Management Systems) which uses natural inputs of precipitation, radiation, and temperature, soil type and management inputs of irrigation, crop type, tillage, fertilizer and pesticides. The outputs include nutrients, metals, pathogens, and sediments in surface runoff and ground-water leachate. This model was applied to model farms of different sizes, animal types and geographic regions. From this model the reductions in pollutant loading of nutrients, metals, pathogens and sediments were calculated for large and medium sized CAFOS that would result from the application of the rule due to nutrient management plans.

How were the benefits valued?

Seven categories of benefits were estimated: water-based recreational use (by far the largest category), reduced numbers of fish kills, increased shellfish harvest, reduced ground water contamination, reduced contamination of animal water supplies, and reduced eutrophication of estuaries. Reductions in fish kills and animal water supply contamination were valued using replacement cost. Increased shell fish harvests were valued using estimated changes in consumer surplus. Water-based recreation was valued using the Carson & Mitchell study described in section 3.3.6 above. Ground water contamination was valued using benefits transfer based on a set of stated preference studies. There was no national estimate of the benefits of reduced eutrophication of estuaries; but there was a case study on one estuary focusing on recreational fishing and using benefits transfer based on revealed preference random utility models.

There are a whole series of potential impacts that were not included in the benefits analysis that would relate to water quality improvements of the rule including human health and ecological impacts of metals, antibiotics, hormones, salts and other pollutants, eutrophication of coastal and estuarine waters due to nitrogen deposition from runoff nutrients and ammonia in the air, reduced exposure to pathogens due to recreational activities, reduced pathogens contamination of drinking water supplies. These impacts were not monetized mainly because of a lack of analytical tools to perform both the impacts and the monetization. Then there are a whole series of ecosystem impacts that were not considered—e.g. the potential changes to aquatic ecosystem functioning that relate to their capacity to produce goods of value to society.

Text Box 3: The Prospective Benefits of the Clean Air Act Amendments

The first Prospective Benefit-Cost Analysis mandated by the 1990 Clean Air Act (CAA) Amendments included estimates of the ecological benefits of reductions in air pollutants to be expected from the 1990 Clean Air Act Amendments (US EPA, 1999). The Agency identified the following potential ecological effects of atmospheric pollutants based on a review of the peer-reviewed literature (US EPA, 1999, p. E-2-E-9):

Pollutant	Acute Effects	Long-term Effects

Acidic deposition	Direct toxic effects toplant leaves and aquatic organisms.	Progressive deterioration of soil quality. Chronic acidification of surface waters.
Nitrogen deposition		Saturation of terrestrial ecosystems with nitrogen. Progressive enrichment of coastal estuaries.
Mercury, dioxins	Direct toxic effects to Animals	Persistence in biogeochemical cycles and Accumulation in the food chain.
Ozone	Direct toxic effects to plant leaves.	Alterations of ecosystem wide patterns of energy flow and nutrient cycling.

The Agency used two criteria to narrow the scope of work for quantification of impacts:

- The endpoint must be an identifiable service flow
- A defensible link must exist between changes in air pollution emissions and the quality or quantity of the ecological service flow, and
- quantitative economic models must be available to monetize these damages.

The Agency provided estimates of three categories of ecological benefits based on standard economic models and methods:

- benefits to commercial agricultural associated with reductions in ozone,
- benefits to commercial forestry associated with reductions in ozone,
- benefits to recreational anglers in the Adirondacks lakes region due to reductions in acidic deposition.

For agriculture, the Agency used crop yield loss functions from the National Crop Loss Assessment Network to estimate changes in yields. These yield effects were than fed into a model of national markets for agricultural crops (AGSIM) to estimate changes in consumers' and producers' surplus. The Agency did not quantify or monetize effects on ornamental plantings, nurseries, or flower growers.

For commercial forestry, the PmET-II model was used to estimate the effects of elevated ambient ozone on timber growth. The PmET-II model is a monthly time step canopy to stand level model of forest carbon and water balances based on maximum net photosynthesis as a function of foliar nitrogen content. The model relates ozone-induced reductions in net photosynthesis to cumulative ozone uptake. Analysis of welfare effects used the USDA Forest Service Timber Assessment Market Model to translate the increased tree growth from a reduction in ozone to an increase in the supply of harvested forest products and computed the changes in economic surplus (consumers plus producer surplus) based on the associated price changes. Because of the lack of data and relevant ecological models, the Agency did not quantify or monetize aesthetic effects, energy flows, nutrient cycles or species composition in either commercial or non-commercial forests.

For estimating the recreational benefits of reducing acid deposition in Adirondacks lakes, the Agency used a published study of recreational angling choices of households in New York, New Hampshire, Maine, and Vermont (Montgomery and Needelman, 1997). This was a random utility model of site choice.
Measured pH of lakes was used as an indicator of the level of ecological services from each lake. The literature on the economics of recreational angling shows that likelihood of success as measured by numbers of fish caught is a major determinant of demand for recreational angling (see Phaneuf and Smith [2005] and Freeman [1995] for reviews). To the extent that populations of target species are correlated with pH levels, pH will be a

satisfactory proxy for fish populations and angling success rates. No attempt was made to identify and quantify other ecosystem services of water bodies likely to be affected by acid deposition.

Modeled reductions in acidification were used as an input to the Montgomery-Needelman (1997) site choice model to simulate the effect of reduced acidification on angler choice and angler welfare. This simulation requires access to the data sued to estimate the model because the benefit measures to anglers depend on individual anglers' travel costs and site alternatives. The Agency's report provides only aggregate benefit measures and does not provide details about how the model and data were used. As a result it is impossible to judge how these transitions were represented.

The Agency also presented an estimate of the benefits of reducing nitrogen deposition in coastal estuaries along the east coast of the US. These estimates were based on the replacement cost or avoided cost method. Reductions in nitrogen deposition reaching Long Island Sound, Chesapeake Bay, and Tampa Bay were estimated. The assumed avoided costs were the costs of achieving equivalent reductions in nitrogen reaching these water bodies through control of water discharges of nitrogen from point sources in these watersheds. As noted in the Methods section of this report, avoided cost is a valid measure of economic benefits only under certain conditions, including a showing that the alternative whose costs are the basis of the estimate would actually be undertaken in the absence of the environmental policy being evaluated, that is, that the alternative's costs would actually be avoided. Since it was not possible to make this showing in the case of controlling nitrogen deposition, the Agency chose not to include the avoided cost benefits in its primary estimate of benefits, but only to show them as an illustrative calculation.

In order to estimate the benefits of reduced nitrogen deposition in coastal estuaries, it is necessary to carry out the following steps:

1. Estimate the changes in nitrogen deposition. The Agency was able to do this for the three estuaries covered in the Prospective Analysis.

2. Use appropriate ecological models to estimate the changes in the populations of species of concern to people. These species include:

   a. Fish and shellfish species that are targets of commercial exploitation;

   b. Fish species that are targets of recreational anglers

   c. Perhaps other species that are of concern to people such as birds and marine mammals.

It might be necessary to estimate population changes not only for species that are resident in and exploited within the estuaries but also for species that use the estuaries for reproduction and shelter of young or that are dependent on species from these estuaries as a food source at some stage in their life cycle.

3. Estimate people's willingness to pay for increases in the services provided by these species. There are models that can be used to do this for commercial and recreational fisheries. But there is very little data on willingness to pay for other types of services such as bird watching and whale watching.

For example, decreasing atmospheric deposition of nitrogen was expected to reduce the deterioration of breeding grounds for fisheries and reduce the habitat loss for aquatic and avian biota. However the Agency was unable to establish the necessary ecological linkages to quantify these fishery effects. Hence it resorted to the avoided cost measure.

6. VALUATION FOR SITE-SPECIFIC DECISIONS

6.1. Introduction

Among the numerous environmental management processes and related decisions that face the U.S. Environmental Protection Agency (the Agency), many are related to specific operating or formerly operated industrial and or municipal sites. The social and ecological implications of such decisions generally are local in nature and affect society at the level of towns, townships and counties rather then at the level of states or regional geographies. Therefore, the goals and performance objectives for these decision processes and their specific decisions need to rely on valuation approaches that are geared to similar levels of spatial sensitivities and are robust enough to adapt to the range of local stakeholder interests that may come to focus through the decision process.

In general the types of regulatory processes that occur at this geographic level under the Agency's or its delegated authority's (i.e. individual states) include a) permits (air, water and waste); b) policies that influence the boundaries for establishing permits (e.g. impaired water bodies designations) and c) administrative orders related to environmental contamination linked to recent non-permitted releases or historical practices prior to current regulatory standards.

In this section we have focused on the regulatory processes associated with the remediation and redevelopment of historically contaminated sites. In particular, we focused on the Superfund program and its efforts to assess the benefits to Ecological services from site redevelopment efforts (Davis, 2001; Wilson, 2005). But ultimately the discussion that follows is generally applicable to any remediation and redevelopment processes for contaminated properties that contain the following basic and common elements:

a) Site Selection - Identification, selection and prioritization of sites
b) Site Characterization - Establish site condition
c) Site assessment – Evaluation of risks and impacts
d) Remedy Selection - Remedial and redevelopment
e) Performance Assessment - Clean-up and redevelopment

f) Public Communication: Assessment results; proposed actions and outcomes

Our goal in this exercise was to explore how the use of valuation methods can positively influence individual steps in a remediation and redevelopment process and lead to a better outcome. As appropriate, individual valuation approaches or methods which could be relevant to specific steps are identified and discussed briefly. In no way is this an exhaustive list of what could be done and the exclusion of a particular method is not implied to mean it is not appropriate for any of the steps discussed.

To explore the opportunities for valuation we have selected to align our analysis with the recent efforts by the Agency's Superfund Program. As noted above, Wilson (2005) provides an assessment of the improvement in ecological service and implied ecological value from the remediation and redevelopment of superfund sites. Although the Wilson paper doesn't actually perform a formal valuation for any of the individual redeveloped properties, it does provide a useful platform from which we can further explore the utility of valuation methods in the remediation and redevelopment process. In preparation for his analysis Wilson (2005) reviewed ~ 40 superfund cases before selecting three case studies which represent urban (Charles George Landfill); suburban (Avtex Fibers) and exurban (Leviathan Mine) environments. We have chosen to analyze and rely on these same three cases to illustrate our discussions about the utility of valuation in the various stages of the remediation and redevelopment process. In addition we have introduced an additional urban example, the Dupage Landfill because it provides a useful counterpoint to the Charles George Landfill example. The Dupage example (http://www.epa.gov/superfund/programs/recycle/impacts/pdfs/dupage.pdf) shows how a early focus on ecological services can provide higher ecological service to the surrounding communities. A brief overview of each of these cases is provided in text boxes 4-7 integrated in the following text.

6.2. Opportunities for using valuation to inform contaminated property decisions

The U.S. EPA Science Advisory Board staff with assistance from the Agency's National Regional Science Council surveyed the regional offices to assess their need for and/or use of valuation information related to Agency regulatory programs. For waste

management and remediation programs (Superfund/RCRA/Brownfield/UST) seven of the eight regions responding indicated that information to help value the protection of ecosystems was needed. Our goal is to help direct the Agency in building the capacity to satisfy that stated need.

The Superfund process and its individual steps or stages are well defined (U.S. EPA CERCLA Education Center, 2005). The steps in the process are provided in Table 5: Steps in the Superfund Process.

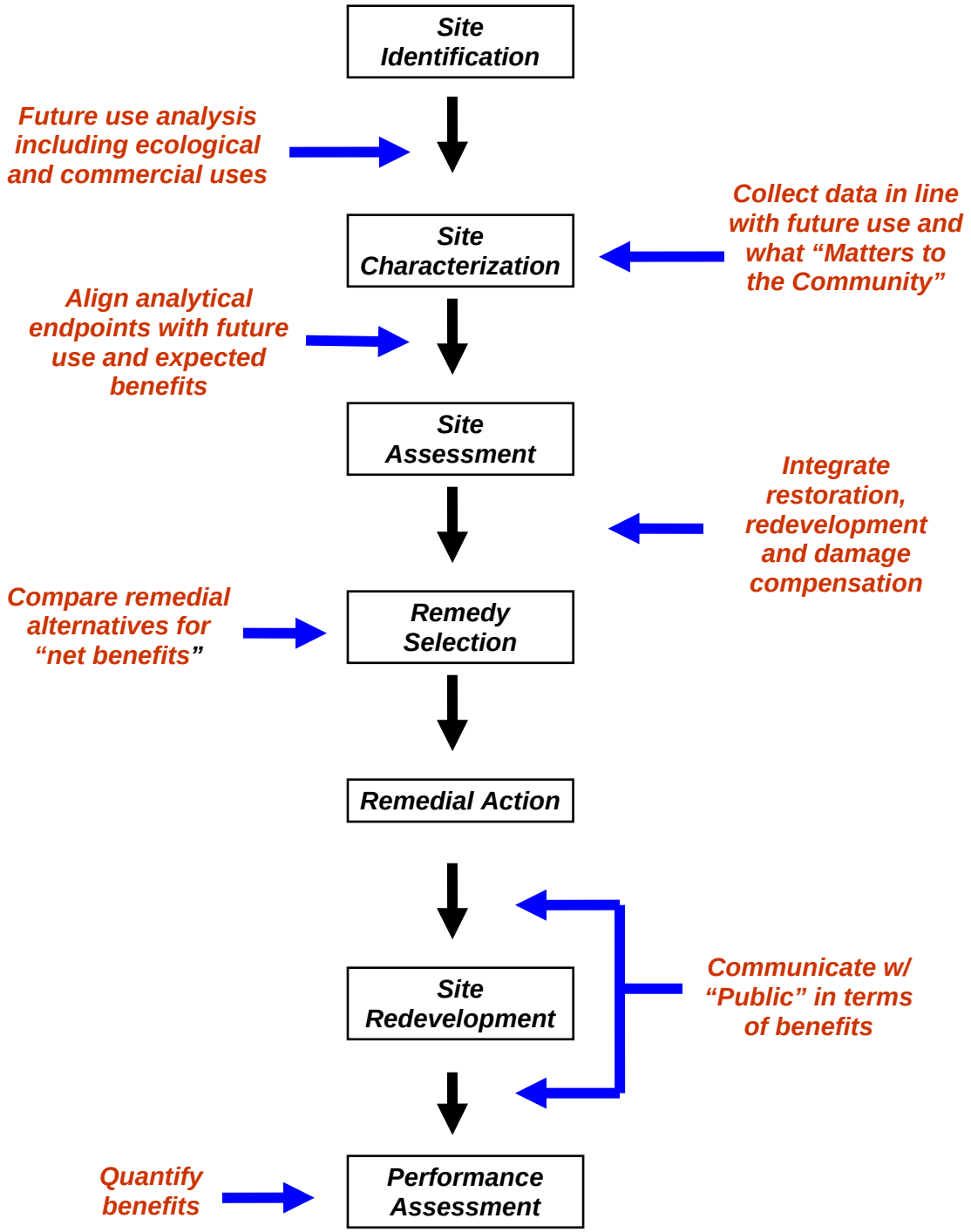
Table 5: Steps in the Superfund Process

1	Discovery and Notification	Initiation: The Superfund cleanup process begins with site discovery or notification to EPA of possible releases of hazardous substances. Sites are discovered by various parties, including citizens, State agencies, and EPA Regional offices. Once discovered, sites are entered into the Comprehensive Environmental Response, Compensation, and Liability Information System (CERCLIS), EPA's computerized inventory of potential hazardous substance release sites (view CERCLIS Hazardous Waste Sites). EPA then evaluates the potential for a release of hazardous substances from the site through these steps in the Superfund cleanup process:
2	Assessment	Preliminary Assessment/Site Inspection (PA/SI) — investigations of site conditions
3		Hazard Ranking System Scoring — screening mechanism used to place sites on the National Priorities List
4		NPL Site Listing Process — list of the most serious sites identified for possible long-term cleanup
5		Remedial Investigation/Feasibility Study (RI/FS) — determines the nature and extent of contamination
6	Decision	Records of Decision (ROD) — explains which cleanup alternatives will be used at NPL sites
7		Remedial Design/Remedial Action (RD/RA) — preparation and implementation of plans and specifications for applying site remedies
8	Cleanup	
9	Closeout	Construction Completion — identifies completion of cleanup activities
10		Post Construction Completion — ensures that Superfund response actions provide for the long-term protection of human health and the environment. Included here are Long-Term Response Actions (LTRA), Operation and Maintenance, Institutional Controls, Five-Year Reviews, Remedy Optimization, and NPL Deletion

More generally, Superfund and related remediation processes are focused on first defining a problem, then characterizing and assessing its potential and actual human health and environmental impacts and finally developing and executing a technical strategy to alleviate or avoid those impacts. M ore recently the evolution of Brownfield initiatives (insert a citation or two) has advanced the integration of a redevelopment focus upstream in the remediation process. In response, the Agency built the Reuse Assessment tool (Davis, 2001) to integrate a focus on Land-Use into CERCLA/Superfund process. By driving remediation and redevelopment closer as a unified process the need to bring value concepts and considerations to the front-end of the process and carry them through the individual steps or stages of the process becomes evident. Net Environmental Benefit Assessment (Efroymson et al 2004) is a recent advance in thinking that provides a framework for using valuation tools to inform the comparison of alternative remedial strategies. Similar efforts are needed for other steps in the Remediation and Development process.

As noted above a generic process that encompasses the remediation and redevelopment would include a series of steps or discrete activities. Figure 6 represents a generic remedial process on which opportunities to include valuation concepts and assessment methods have been identified. As is clearly shown, early recognition of future uses and ecological services that matter to people will carry through to inform assessment of the site and the ultimate selection of remedial actions and redevelopment options. Optimally, by expressing expected and/or capture benefits will lead to more effective communication with concerned publics. The opportunities and utility of such adaptation of valuation methods to this new merged process is discuss in the following sections.

Figure 6: Changing Focus from Remediation to Redevelopment Would Benefit from Increased Integration of Valuation Analysis with Traditional Process Steps


Site Identification
Future use analysis including ecological and commercial uses
Site Characterization
Collect data in line with future use and what "Matters to the Community"
Align analytical endpoints with future use and expected benefits
Site Assessment
Integrate restoration, redevelopment and damage compensation
Compare remedial alternatives for "net benefits"
Remedy Selection
Remedial Action
Site Redevelopment
Communicate w/ "Public" in terms of benefits
Quantify benefits
Performance Assessment

1

In general, valuation methodologies should be most useful for identifying how a site and the current or potential ecological service flows matter to the surrounding community. Such methods should be focused on determining what benefits can be or have been derived from the site and how any potential effects on the ecological components diminish those benefits. When the ecological services that matter to people are well defined and when the assessments of ecological production and risk can be coupled to these specific services, then the outcome is likely to be a remediation and redevelopment plan that is targeted on what really matters to the local community. Therefore a key recommendation is that consideration of ecosystem services and their benefits to human well-being and other forms of value need to be considered from the earliest stages of addressing contaminated properties.

Even as early in the management process as site selection or prioritization, tools which allow for comparison among sites for their ecological services potential could be informative. Additionally, valuation can be used to capture the benefits linked to site ecological attributes and identified ecological services to the surrounding community. Data that supports or aids in the design of benefits assessment should be considered in the design of any site characterization plan. While a typical site characterization is focused on the aerial extent of chemicals and their range of concentration in site media (e.g. Ground and surface water, soil and biological tissue), a plan that also collects information to define ecological services flows and how they matter would lead to a better alignment of ecological risk and economic benefit assessments. Aligning risk and benefits assessment should be a critical objective for the Agency as it will assure that the remedial actions selected for consideration will address the restoration of the benefits derived from any important ecological service flows that have been diminished or disrupted. As well, aligning risk assessment endpoints with ecological services and the derived benefits from those services should lead to improved a) alignment with community goals; b) ability to better perform meaningful benefits assessment and c) ability to communicate proposed actions and d) ability to monitor and demonstrate performance

As has been pointed out through the introduction of Net Environmental Benefit Assessment (Efryomson et. al., 2004) valuation can be a useful approach to aid selection among remedial technology options by weighing and comparing the benefits among the

options. Incorporating valuation methods into the NEBA framework would provide the basis for balancing trade-offs between risks and benefits of the ultimate remedial design. Additionally it will aid in keeping the set of ecological services preferred by the community as driving function in the prioritization and selection of remedial and redevelopment actions.

Ultimately, the test of the process is to what degrees were the ecological services and associated benefits of importance to community either protected or restored. If as originally recommended, values have been broadly explored and effectively highlighted and integrated into the site assessment and remedy selection processes then measures of performance will be apparent. Ecological measures of productivity or aerial extent of condition which are directly linked in an understandable manner to valued ecological service flows will be useful in tracking the performance of remediation and redevelopment processes. Advancing the Agency's capability to do performance evaluation both in real time and retrospectively will help the Agency better justify in the future the overall performance of remediation and redevelopment of contaminated sites.

Finally, the remediation and redevelopment of a property is really an exercise in social engineering that encompasses more then just the biological, chemical and physical sciences and engineering principles that historically have underpinned the remediation process. Therefore, effective communication with stakeholders participating in the management process and those outside is a critical element to success of the management process. Both of these audiences will be bringing a value set to the table when they are assessing any proposed actions or evaluating the results of any action taken. Therefore having a strong alignment between the ecological services valued by these audiences and the expected or actual outcomes will facilitate effective communication.

6.3. Use of source examples to illustrate recommendations

In Part One, Section 6 of this report, a series of high level recommendations were provided. In essence, it was recommended that ecological values and benefits derived from ecosystem services should be considered from the outset when framing any analytical process to support Agency decisions and associated actions. The recommendations direct the Agency to broaden its consideration of the types of ecological values and to align them with what matters most to the people involved or

affected by the decision. Additionally the Agency is encouraged to explore expanded use of economic and ecological models to characterize and measure the values associated with environmental change.

In order to facilitate the charge to expand its focus on values, it is recommended that from the outset that expertise and opinions be brought to the process by integrating technical disciplines and engaging interested and affected stakeholders. Ultimately by aligning those ecological services that benefits people the most with ecological production functions that drive their availability, the Agency will be able to focus its actions to produce maximum protection or in the case of contaminated site maximize restoration of benefits.

In the following text we have taken and adapted those general recommendations to the site-specific application context. In addition we have supported these site-specific recommendations with lessons gleaned from a series of Superfund examples at the Urban (Charles George and Dupage Landfills), Suburban (Avtex Fibers) and Exurban (Leviathan Mine) demographies. Text boxes 4 and 5 provide background on the urban landfill cases. Text box 6 and 7 provide background on the suburban and exurban cases respectively.

6.4. Recommendations

The recommendations outlined in Part 1, Section 6 of this document provide general guidance to the agency on where to focus their energies to improve the use of valuation methods in their programmatic and policy activities. The following recommendations for site-specific or local management decisions, build on that general guidance. An overarching premise for the following recommendations is that early and strategic insertion of valuation methodologies will maximize their influence on any local or site-specific decision process.

<u>Recommendation #1</u>: At the beginning of the process, broadly define the range of ecological services and associated value(s) recognized as important by key stakeholders and the community at large as attributable to the site or locale. To achieve this objective:

a) Explore the utility of a variety of group process (e.g. Deliberative facilitated) and survey methods (e.g. Social-Psychological or "attitude") to engage stakeholders in this process from the outset.

b) Consider the many sources of ecological value including both instrumental and intrinsic.

c) Consider not only current or diminished ecological services, but also the potential for developing or enhancing ecological services not presently utilized.

Recommendation #2: Appropriately involve the right mix of interdisciplinary collaboration from physical, chemical, biological (ecology, toxicology etc.) and social scientists (economists, social psychologists, anthropologists, etc.) in line with site-specific considerations and conditions and the specific step in the process.

Recommendation #3: Clearly demonstrate the alignment between ecological services the ecological functions that produce those services and potential positive or negative effects from current conditions or proposed agency actions. To achieve this objective:

a) Develop the capacity to utilize an ecological – economic conceptual model to inform the site assessment design.

b) Develop the "accounting rules" to recognize and avoid double- counting or under-counting the benefits from ecological service flows.

c) Develop approaches (e.g. Multi-Attribute Utility Analysis, Bayesian) to sort, weight or otherwise prioritize ecological services for primacy for actions and also to weigh benefits derived.

Recommendation #4: Expand the variety of methods in the Agency's arsenal to quantify the ecological service, to describe ecological production functions and to capture in monetary and non-monetary terms the value lost or gained from current conditions or some proposed agency action.

a) Explore the current state and extent of ecological production function models

b) Develop a strategy for adapting existing general models to site-specific applications

Recommendation #5: Develop the capability to utilize valuation techniques to provide a basis to demonstrate Agency performance and communicate the expected or actual outcome from Agency actions.

Recommendation #6: Create formal systems and processes to foster an information sharing environment.

a) Actively document lessons-learned from applications of valuation methods and share broadly among program and project managers.

6.5. Source Example Analysis

Recommendation #1. At the beginning of the process, broadly define the range of ecological services and associated value(s) recognized as important by key stakeholders and the community at large as attributable to the site or local.

Broadly define ecological services early in process. The urban examples of the Charles George (See text box 4) and the Dupage County landfills (See text box 5) strongly show the difference in outcome that can be produced by engaging with community to focus on the ecosystem services of importance to them. Although there was no evidence of formal valuation methods at the onset of either example, the focus on how the site will provide future benefits and the inclusion of additional disciplines form lead to a more positive outcome for the Dupage county community.

At the Charles George landfill, ecological values or future uses were not considered at the start. The human health risks at this site were so salient at the time that they were discovered that they controlled the focus of the subsequent decisions. When the landfill site was capped and the water system from the city of Lowell, MA was extended to the affected community, the health and safety concerns were addressed. Although an effort to make the site work environmentally has now begun (insert URL for

restoration plan), still some 20 years later, the potential for ecological services remains untapped.

By contrast, the remediation and redevelopment of the Dupage County Landfill site, now known as the Blackwell Forest Preserve, appears to have been motivated largely by the need to address intrinsic (rare birds; e.g., hawks) and recreational (e.g., hiking, bird watching, boating, camping, picnicking, sledding, etc.) benefits. The remediation effort succeeded. Listed as a Superfund site in 1990, "a once dangerous area is now a community treasure, where visitors picnic, hike, camp, and take boat rides on the lake." http://www.epa.gov/superfund/programs/recycle/impacts/pdfs/dupage.pdf

Engage key stakeholders. Public input for the Axtex Fibers case was an evolving process of growing complaints about offensive sights and smells and about contamination of drinking water wells. Over several decades, local government and environmental protection agencies made tests, filed thousands of complaints and took various regulatory actions that ultimately resulted in the listing and designation as a Super Fund Site. Once the site was listed and a management process established there was a clear effort to engage stakeholders through the Multi-stakeholder process for development of the master plan. But as in the case of the Dupage landfill, although ecological services or at least ecological components were considered, it is not clear there was any systematic assessment of "what people cared about" regarding the Avtex site. Whether a more formal assessment of values would have reached a different or clearer description of community values is an open question. In any case, the commissioned Master Plan (insert reference or URL) that was developed in interaction with a "Multi-Stakeholder Group" implies that ecological restoration and ecosystem services (especially relevant to water quality) were important considerations for cleanup and redevelopment of the site. A substantial part of the site plan is devoted to restoration of forests consistent with natural conditions at the site, and waste pits are being redeveloped as ponds and meadow/wetland areas to provide important runoff control, water purification and wildlife habitat services. Much of the redevelopment of the site is directed at enhancing aesthetic values by restoring naturalistic landscapes to be enjoyed by recreational users, nearby residents and passing tourists.

Define the ecological services that matter to people, Determining what people care about requires a carefully constructed and systematically implemented program integrating assessments of multiple values using multiple methods to fairly and faithfully reflect the perspectives of multiple stakeholders. There is no simple recipe for accomplishing this, and no simple algorithm for calculating values and summing them up to make a decision. Value assessments serve to support decisions that must in the end be based on the judgment of administrators charged by society with that responsibility.

The Leviathan mine is a good example of how the Agency is often faced with the need to consider a complex array of competing interests. In this case the Agency is faced with a clear dichotomy between the ecological services valued by the full time resident native people and the community of occasional recreational user. The recreational users would gain from the cultural services associated with hiking, fishing and camping. While the Washoe tribe which lives in the area year round would benefit from the resource both as a provisioning service for food but also from the spiritual and cultural services.

Additionally the Leviathan mine case study highlights the need to consider the existence or intrinsic value of the ecosystem. For example, the ecosystem near the Leviathan mine site provides habitat for threatened species such as the Lahontan cutthroat trout and bald eagle, which many tribal and non-tribal individuals might value even though they provide no direct instrumental or use value. In considering site restoration or remediation, or measuring damages from contamination at the sight, the Agency would be missing the primary sources of value if it limited consideration to standard types of use value and did not consider these other sources of value as well.

Finding effective ways to both understand the values of disparate users for the same resource and to effectively weigh their interests in restoration and redevelopment of the site is not a small challenge. To include the relevant sources of value in an assessment, the Agency has to determine what aspects of the ecosystem generate these values, i.e., what aspects of the site contamination are of greatest concern to people. For the Leviathan Mine case, it is likely that this would have to be considered separately for tribal and non-tribal individuals, since the sources of value are likely to be different for these two groups.

Recommendation #2. Appropriately involve the right mix of interdisciplinary collaboration from physical, chemical, biological (ecology, toxicology etc.) and social scientists (economists, social psychologists, anthropologists, etc.) in line with site-specific considerations and conditions and the specific step in the process.

Integrate disciplines. Interrelationships among experts and between experts and the affected publics form a key component of any hazardous site assessment, planning and implementation program. Ideally, collaborations among all relevant experts and communications with affected publics/stakeholders begins very early in planning and decision making and remains active throughout implementation and post-project monitoring and evaluation. A key point for collaboration among expert disciplines is in the development of alternative management scenarios, particularly translating physical and biological conditions and changes at the site into value-relevant outcomes that can be communicated to stakeholders.

The Leviathan mine case provides another instance of the need for integrating unique or non-traditional disciplines into efforts to understand what affected human population's value. Because of the unique cultural and spiritual values associated with ecological services, anthropologists could be involved in understanding and quantifying or characterizing the value of the ecosystem services to the Washoe Tribe. Likewise, in order for economists or others to try to estimate existence value for an impacted species (e.g., fish), it is necessary for them to work closely with ecologists to determine the likely impact of any change (or proposed project) on that species (e.g., effect on fish population) so that this change can be valued.

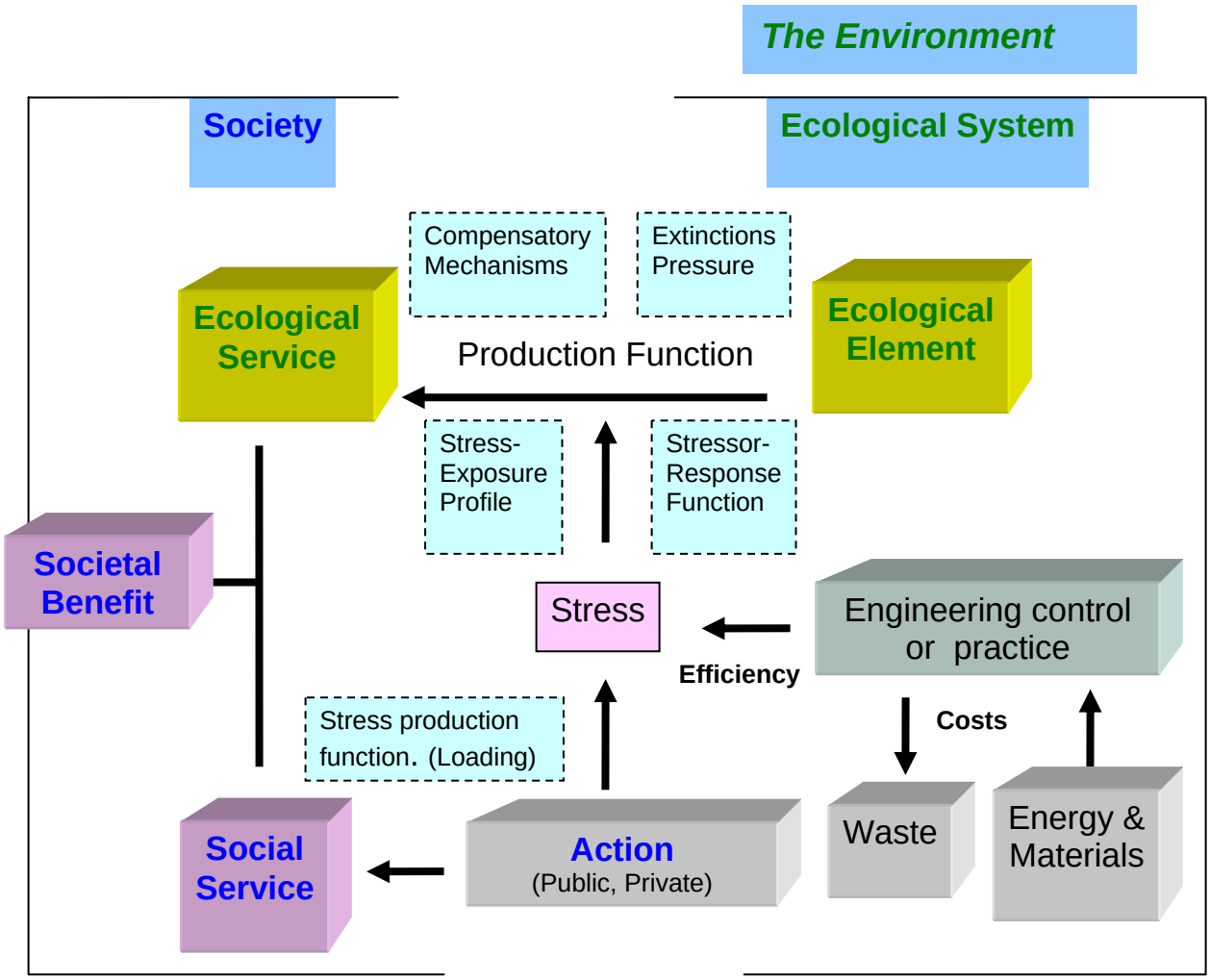
Recommendation #3. Clearly demonstrate the alignment between ecological services, the ecological functions that produce them and potential positive or negative effects from current conditions or proposed agency actions.

The call for alignment between ecological services, production and risks in recommendation # 3 is at the technical core of performing any of the risk or benefit assessments associated with the remediation and redevelopment of contaminated property. Unfortunately none of the source examples chosen for this example provide a demonstration of active intention to create such alignment. For the most part the best we

can do is use the examples to illustrate where for those cases we believe such alignment would have influenced the results in a positive manner.

Utilize an ecological-economic conceptual model. Developing a conceptual model is an expected and standard practice in performing ecological risk assessments for contaminated site evaluations. A conceptual model which integrated and aligned the ecological aspects of risk with economic benefits from existing or foregone ecological services would facilitate better alignment between remediation and redevelopment. Figure 7 provides a generic representation of the linkage between the environmental stress generated by public and private efforts to generate societal services, its potential (i.e. risk) for affecting ecological production and their associated ecological services to society

Figure 7: Ecological Services Linkages Conceptual Model



The potential benefit to the Agency from developing the capacity to use conceptual models that integrated ecological and economic attributes of the site is highlighted by the Avtex Fiber case. Health threats to workers and to nearby residents were highly salient concerns and strongly guided initial management plans and actions at the Avtex site, potentially reducing opportunities to recognize and address important ecosystem risks and associated ecological services. Technical risk assessors and public observers/participants would have benefited from a clear and comprehensive model of the ecological roles being played, and that potentially could be played by the Avtex site.

Early concerns about contamination of groundwater and discharge of toxins into the Shenandoah River focused attention on water quality.

A noteworthy feature of the Avtex Fiber process was the development of a Master Plan. There is evidence form that plan that ecological services were considered. For example, aquatic basins constructed to contain contaminants on site were also designed to restore important ecosystem services as well, including providing safe habitat for water fowl, runoff control and water purification services. In this regard some at least rudimentary ecological production function is implied by the plan although not documented or for that matter their benefit quantified. It is not clear that other aspects of the Avtex Fibers site Master Plan were as effective at addressing ecological risks or at capitalizing on opportunities to enhance ecosystem services in the redevelopment of the site.

The development of an ecological-economic conceptual model would have systematically informed greater integration of building ecological service into remedial design and future uses. To that point, recreational and aesthetic services were clearly important considerations for many features of the plan, but it is not clear whether the specific allocation of facilities and uses to spaces within the site was guided by any comprehensive ecological model. For example, it is not clear whether the particular pattern of restored forests and wetlands, developed recreation areas and industrial park produces the best possible outcomes for protecting ecosystems and ecosystems services. It is possible that different siting and design of the soccer fields, for an example, might have returned the same recreational benefits while achieving greater ecosystem services in the form of wildlife habitat, water quality or aesthetic values for visitors and/or nearby residents. The declared ecological, “green” focus of the industrial park component of the master plan implies that ecological concerns will be paramount in the selection of industrial tenants and in the siting and design of facilities, but no ecological model for achieving this goal, or monitoring progress toward it is presented. This leaves open the prospect that future industrial, recreational and tourist developments and uses at the Avtex site might simply substitute one set damages to ecosystems and ecosystem services for another.

Need for "accounting" rules to count benefits. Ecosystems and their numerous components are linked in an intricate and complex network of biological, chemical and energy flows. By looking at impacts to individual organisms or components and their associated services in isolation, there is a serious opportunity for double counting service losses and or benefits generated by Agency actions.

For example, the listing of services (aquatic biota and habitat, riparian vegetation, terrestrial wildlife, recreational uses, and tribal uses) in the Leviathan Mine case does not seem to be very useful for sorting out the different things to be valued. It does not identify mutually exclusive services and seems to have a high likelihood of double counting. It also does not seem to adequately distinguish between "inputs" and "outputs". As well, the question of why we care about protecting habitat or riparian vegetation is not clearly addressed. Is it because we care about the populations it supports for their own sake, or because these populations are an input into something else we value, such as recreation? Take insect populations, if we care about the insects for their own sake, then maybe this should be included as an existence or intrinsic value. If we care about them because they are a food source for fish and we care about fish, then we should value the change in fish brought about by the change in insects but not value both separately, i.e., we should view both clean water and insects as inputs into the production of more fish, and value either the inputs or the output. Of course, then there is the question of why we value the fish because of their existence, their recreational use, or their cultural significance to the Washoe tribe. Perhaps part of this whole exercise is to first try to answer the question of why we value the insects or fish. It seems we need to know this before we can figure out how to measure how much we value them.

Similarly, the listing of services by Wilson (2004) shown in Table 6: Ecosystem Service Matrix for Leviathon Mine (from Wilson, 2004) based on the U.N. Ecosystem Millennium Assessment (2005) definitions of ecological services is not very useful for valuation purposes, and in some cases we believe it could create confusion in valuation. For example, it is not clear how or where the use of surface water or groundwater for drinking would fit in Wilson's list. Is the service from "Freshwater Regulation" intended to include drinking water or is it intended as an input into aquatic and other habitat-

related services? The valuation approach used is likely to be different depending on which of these services freshwater regulation is intended to reflect.

Table 6: Ecosystem Service Matrix for Leviathon Mine (from Wilson, 2004)

Ecosystem Function	Ecosystem Service
Regulating	Disturbance Moderation • Flood prevention from on-site evaporation ponds • Regulation of surface water runoff and river discharge during snowmelt and heavy rain events
	Freshwater Regulation • Restoration of groundwater discharge beneth the pit and waste-ore piles • Non-hazardous surface water drainage into Leviathan Creek, Bryant Creek and East Fork River
	Wildlife Habitat • Nursery, feeding and breeding ground for indigenous fish speies including the threatened Lahontan cutthroat trout • Restoration of habitat and feeding habitat for the threatened Bald Eagle • Maintenance of riparian vegetation habitat for mammals, birds , amphibians and insects
Supporting	Soil Formation • Restoration of productive floodplain soils in the leviathan-Bryant Creek watershed and the East Fork of the Carson River
Provisioning	Food and Raw Materials • Edible freshwater fish • Pine nut harvesting by Washoe tribe
	Ornamental Resources • Raw material for traditional Washoe Tribal crafts
Cultural	Recreation and Amenity • Improved hiking and camping opportunities • Recreational fishing
	Inspirational and historic • Washoe Tribal heritage site • Spiritual and ritual uses such as spiritual bathing, and cleaning religious implements

Perhaps a better delineation of services (defined as outputs rather than inputs) would be the following:

a) Water used by Washoe Tribe members and others for washing and drinking

b) "Existence" or intrinsic values (broadly defined, based on moral or other principles) from threatened and other species (e.g., cutthroat trout, bald eagles, and other impacted species of concern)

c) Non-consumptive use values of wildlife (e.g., people like to view bald eagles and other species)

d) Harvesting (hunting, nuts, fish) by Washoe tribal members

e) Cultural/spiritual and ceremonial value of land used by Washoe tribal members

f) Water flow regulation (e.g., reduction in flooding from snowmelt or runoff)

g) Non-tribal recreational services (e.g., fishing, hiking, camping)

h) Value of the natural process leading to ecosystem outputs, above and beyond the value of the outputs themselves (e.g., preference for natural processes over man-made ones, or native species over introduced species)

In any case, it is clear that there is a need to establish some accounting guidance for working with complex social and ecological situations.

<u>Align ecosystem services with ecological production functions and impacts/risks</u>. To achieve that objective of alignment the Agency will need to bring forward in the planning process for site remediation and revitalization a robust discussion of what are the ecological services and to what degree they matter to the affected local community or the ability of the environment to sustain its integrity. To some degree the Agency has already settled on the concept that ecological risk assessments need to be built on an ecological construct and a conceptual model that is linked to an assessment endpoint. The gap in practice maybe as simple as doing a more thorough analysis of the breadth of ecological services and how they matter to people. This will present technical challenges as today the design of ecological risk assessments are dominated by what toxicological data we have in the literature for a limited range of species. It is very likely that the species data we have will not link well to the ecological services that matter. This may require the agency to revisit its assessment approach for chemical exposures from an ecological services perspective rather than toxic response of individuals. In the mean

time, more attention to creating the alignment between ecological services, the assessment and measurement endpoints used in the risk assessment and the ability of economists to place a value estimate will likely lead to significant improved outcomes in efforts to revitalize land. In addition, a significant Agency effort to estimate the population or community level consequences of chemical exposures on ecological service flows will advance this objective greatly. To do that the Agency will need to develop the capacity to adapt and apply ecological production models in its contaminated sites assessment processes. These models are the real bridge between risk estimates and subsequent injury or damage projections and provide a major piece of the puzzle to quantify and value the impacts of chemical exposures as well as the remedial and restoration alternatives.

Although other trustee Agencies NOAA and USF&W are the regulatory leads for Natural Resource Damage Assessment (NRDA), the ecological risk assessments and conceptual models produced by the Agency in the remediation process are often the basis for damage assessment. That extrapolation from risk to injury and then onto damages is often a significant point of departure in the dialogue for the Agency and their trustee partners with the parties responsible for the damages. The uncertainty in estimates of chemical exposure, toxic response and therefore the estimate of risk makes using these data as a surrogate for injury to the environment controversial and therefore the resultant damage claim for reduction in human use or ecological services is likely to be challenged. Damages are an expression of the needed restitution for lost or forgone use of ecological services. To link risk or potential for injury with actual loss of service and the estimate of the values of that service (i.e. damages), will require linking ecological services with the environmental components producing those services and then defining the risk to or likely response of those ecological components to chemical exposures.

The Leviathan mine case illustrates both how the concept of ecosystem services has and can be used in damage assessment and restoration, as well as some of the challenges associates with delineating services in a way that is useful for valuation. One could suggest that if the agency can achieve the recommendation to align ecological services, their production functions and risk profiles then it would also benefit the ability

of resource trustees to appropriately assess injury, define restoration goals and calculate damages

In the Leviathan Mine example, impact or injury is defined not only as exceeding of some standard (e.g., water quality or drinking water standards) but also as concentration or duration sufficient to cause a loss of services provided by the resources to the general public in addition to unique service losses to the Washoe Tribe. Thus, the concept of ecosystem services plays a key role in defining or focusing categories of possible injuries to further evaluate.

Similarly, the concept of ecosystem services underlies the use of Habitat Equivalency Analysis (HEA; also know as Resource Equivalency Analysis or REA) to determine compensation for damages. In principle, application of the HEA concept requires a determination of the flow of ecosystem services that would have been provided by a given site "but for" the contamination and a comparison of this flow with the flow of ecosystem services provided as a result of a restoration or other project designed to generate an equivalent service flow. Ideally, the value of the ecosystems services under the two would be equal. In order to apply this concept, it is necessary to delineate and implicitly value the service flows.

How can the impact of the site on these services be estimated? The Leviathan Mine Natural Resource Damage Assessment Plan (NRDAP) gives detailed information on concentrations of key pollutants (in particular, heavy metals such as cadmium, zinc, copper, nickel, and arsenic) in surface water samples, groundwater samples, sediment samples, samples of fish tissues, and insect samples at various distances from the mine site. These concentration levels are compared to levels at reference sites (since historical information is not available) to illustrate impact. In addition, they are compared to baselines determined by water quality criteria (under the CWA), drinking water standards (under SDWA), etc. to quantify the magnitude of the impact/injury. In general, unacceptable risks are defined based on toxicity thresholds or other concentration criteria, as well as on the extent to which species impacts based on comparable concentration levels are documented in the literature.

Once the impacts on water quality, sediments, etc., have been determined, they need to be translated into predicted changes in the flows of the services listed above. In

principle, it requires the estimation of an ecological production function. For example, to see if recreational fishing is likely to be significantly impacted, we would need to estimate the impact of the site on the fish population in the nearby water body. This requires estimation of the impacts of the changes in things like water quality, streambed, bank sediments and riparian vegetation, on fish population, both directly and indirectly through their impact on the insect population. For example, if we know that there are elevated levels of arsenic, copper, zinc, cadmium, etc., in insects and fish tissue, how do we use this information to predict an overall impact on the fish population? In most cases, an ecological model for doing this at a particular site such as the Leviathan mine will not exist.

How then should EPA proceed in trying to look at not only the impact on ecosystem resources but also (or instead) the impact on ecosystem services? At this stage, EPA might instead look at the scientific literature to see what it says about how sensitive the insects and fish species of concern here are to these types of stressors and then ask expert ecologists to provide some expert judgment on the likely magnitude of the impacts in this specific case. (This would be akin to an "ecological impact transfer", similar to the notion of benefits transfer.) In fact, the Leviathan Mine NRDAP suggests this.

In addition, the Leviathan Mine NRDAP suggests looking at, for example, the fish population downstream of the mine and comparing it to the population in a reference location. More generally, it suggests comparing not only fish populations but riparian vegetation, the composition of the benthic community, wildlife populations, etc. near the mine and at a reference site. Such a comparison can determine the damages resulting from the mining activity (which is most useful in an NRDA policy frame), but it cannot directly predict the impact of proposed remedial actions on ecosystem services (which is needed for a policy frame relating to evaluation of remedial actions), unless we make the assumption that the remedial actions will be 100% effective in restoring the ecosystem services to their original level (presumed to be the level at the reference site). Short of this, we would again need to predict the impact of the remedial actions on the ecosystem resources and then translate those into predicted changes in ecosystem services using an

ecological production function. Such a balancing act could be assisted through the use of comparative tools such as Net Environmental Benefit Analysis (Efroymson et. al., 2004).

Recommendation #4. Expand the variety of methods in the Agency's arsenal to quantify the ecological service that matter to people and to capture in monetary and non-monetary terms the value lost or gained from current conditions or some proposed agency action.

Expand methodological capacity. Part 3 of this report provides an overview of a broad range of methods that could be explored for the integration of valuation into the typical contaminated property redevelopment. For any of the source examples selected to highlight local decisions, their decision making processes could have benefited from the application of a number of these methods. The Agency should be exploring the use and/or adaptation of many of the techniques listed in the methods section of this report to: a) engage stakeholders to define what they value; b) help align the sites risk assessments with expected benefits; c) test alternatives strategies for redevelopment to achieve those benefits and d) improve communications of proposed actions and their performance.

Future uses that matter to stakeholders. Determining local stakeholder interests with regards to preferred future property uses and the ecological services derived from that redevelopment scenario is an important starting point. Survey methods or facilitated dialogues would be useful methods to achieve this objective. In helping to frame the dialogue with stakeholders, methods such as the Environmental Benefits indicators (Boyd, 2004; Boyd and Banzhaf, 2006) or the Biodiversity Indicators (Grossman, 2004; Stoms et. al., 2005) may be very suitable for helping Agency's site managers understand the ecological service potential from future uses and provide the basis for a Deliberative Facilitated-Group Process (need citations or section reference) or mediated modeling (need citation or section reference) exercise.

The counterpoint represented by the urban examples show that even the most rudimentary dialogue about future use can lead to an outcome with greater service to the community. At the Dupage Landfill site, it seems that only a qualitative focus on the utility of ecological services lead them to recognize that in a very flat landscape, even a 150-foot hill, if properly capped and planted, would be a welcome refuge for people as well as wildlife. The Dupage Forestry District had a sense of the ecological potential of

the area particularly for hawks – and where hawks abound, so will birders to watch them. In this case, the difference is not one of methodology so much as conception – once planners "see" an area as having ecological potential, it may be a fairly easy matter to point to qualitative differences to show, by way of analogy and example, likely quantifiable or monetizable consequences. It might be a valuable learning useful exercise for the Agency to go back to a case like the Dupage or the Charles George Landfills and develop a valuation assessment plan. Such a plan would create a map of possible methods the Agency and responsible parties could use to integrate valuation into the decision process

For the Avtex Fibers site, deliberative group processes involving stakeholders and relevant experts (including historians) would have provided an effective approach to identifying the ecosystem and ecosystem service values of most concern to stakeholders. Systematic assessments of ecological values and of historic and sense-of-place values (assuming these were identified as important) are not well developed. Stated-preference monetary assessment methods, such as contingent valuation surveys (methods citation or reference to sec #) might be applied. People, however, have been shown to have difficulty expressing consistent willingness-to-pay estimates for such non-commodity outcomes and some people find assignments of dollar values to be ethically offensive in this context. Social-psychological "attitude" survey methods (See section X.X) could provide relative measures of preferences (importance, acceptance) for any defined population of stakeholders for the array of ecological, historic and sense-of-place outcomes across a defined set of cleanup and restoration/redevelopment options for the site.

For the Leviathan example there are at least three ways that information about the impacts of greatest concern to affected individuals might be obtained. The first would be to gather information about the relative importance of the various services in this particular context through focus groups, mental models, mediated modeling, deliberative processes and other similar methods. Similarly, for services relating to Tribal uses, anthropological or ethnographic studies based on detailed interviews can be used to determine the ecosystem features of most importance and the characteristics necessary for suitable restoration or replacement.

The second approach would be to gather some basic information that could be used to judge the importance of different services. This might be of the type used to construct environmental benefit indicators. Examples would be: Water use data for the Washoe tribe and others in the vicinity of the site (e.g., sources, quantities, purposes); harvesting information for the Washoe (e.g., what percent of their harvesting of nuts, fish, etc. comes from the area impacted by the site); recreational use data (Number of people visiting the area of the national forest impacted by the site for hiking, camping, fishing, wildlife viewing); data on flooding potential and what is at risk in the vicinity of the site; data on spiritual/cultural land use practices by the Washoe. The information regarding the Washoe could be collected through interviews. It is not clear whether some of the other data exist or would have to be collected.

The third approach would be a review of related literature and previous studies to draw from what has been learned in other contexts. For example, previous Social Psychological surveys (not specific to this site) or other expressions of environmental preferences/views (e.g., outcomes of referenda, civil court jury awards, citizen juries, etc.) might provide insight into what people are likely to care about in this context. Similarly, previous contingent valuation studies of existence value might provide some (at least partial) indication of the likely importance of impacts on species such as bald eagles (e.g., if they show that existence value is large). Likewise, previous studies of the value of recreational fishing (e.g., from travel cost models) could be coupled with the use data above to provide an initial indication of the importance of the impact on recreational fishing.

<u>Aligning ecological services with risk assessment</u>. There is not a single method that could be identified which is focused on mapping prediction of ecological risk with Production functions and the services that derive from those structural or functional ecological units. Qualitatively or visually the linkages can be represented by the creation of an ecological – economic conceptual model as discussed previously in this section. Once a visual representation of the relationship between a stressor impacting ecological production and the change in an ecological service has been mapped, the agency is still left with the challenge of quantification.

The agency has already approached the development of complex ecological risk assessment modeling tools (TRIM, EXAMS, AQUATOX) to estimate the fate and effects of chemical stresses on the environment and has even coupled such exposure-effects models with ecological production models to estimate population level effects. Although there not many examples of such integration it would not be impossible for the Agency to focus on expanding such capability by exploring the world of existing ecological production and ecosystem level models that exist in the literature (Roughgaren, Joan 1998a and 1998b; Roughgarden, Jonathan, 2001).

A major gap in the current ecological modeling capability is coupling the aforementioned modeling systems with models (or modules) that link ecological production models with explicit ecological services which can be quantified. This is very important because such a tool could be used not only to assess impacts and their acceptability but also as the quantitative basis for looking at the benefits derived by investments in alternative remedial and redevelopment strategies. Without this capability the agency is left with the narrow ability to look at risk reduction as the primary ecological benefit from any action.

Testing remedial and redevelopment alternatives. Currently the typical comparison of remedial alternative strategies includes two tests. The first test being does the action control risk to an acceptable level. All of those technologies that pass that minimal benchmark then go through a second test for cost-effectiveness. So if all technologies are adequate with regard to risk reduction then the least costly is the obvious choice. What such an approach does is decouple remediation and development, which leads to a delayed development process possibly off mark from what matters to key stakeholders.

If alternatives can be compared based on benefits generated then it opens up a number of methods that could be used to compare alternatives with or without stakeholder direct involvement. As mentioned previously, Net Environmental Benefit Analysis (Efroymson, 2004) is a framework for comparing remedial/redevelopment alternatives on a basis of benefits generated. Obviously the units of those benefits could be in either monetized or non-monetized units. For example in superfund sites the value of an action could be expressed through methodologies such as Habitat Equivalency

Analysis (HEA) or Resource Equivalency Analysis (REA) (Section Reference). Although HEA and REA generally produce results in ecological units over time (e.g. discounted service acres years) the cost of creation or replacement of those ecological units can be estimated in monetary terms (i.e. replacement cost) . This approach does not provide a direct measure of the value of ecosystem services, but it does support a comparison of the services provided under different options. For the most part, we are looking to achieve a reasonably precise and representative measure of relative benefits for comparing alternatives. So to the degree that other methods which measure outcomes purely in ecological terms, such as the Biodiversity and Conservation Values approach (Grossman, 2004, Stoms et. al., 2005) provide a useful basis for comparison among options they might also be useful in conjunction with NEBA.

Comparison of alternatives via monetary/economic valuation methods might include hedonic pricing studies to determine the economic impacts of the identified cleanup and redevelopment options on adjacent residential property values. As well, input-output models might be used to compare expected gains to the local economy across the feasible set of redevelopment scenarios. Monetary/economic assessments and models might also be used to estimate the expected long-term contributions to the local economy from industrial development versus recreation/tourism-focused use options.

If stakeholders are involved in testing alternatives then their preferences or weighting of alternatives could be assessed directly through group deliberative value assessment processes. This would allow non-monetary methods such as ecological value assessment methods to be used as a basis to compare changes in biodiversity, habitat quality, energy flow and other indicators of identified and accepted bio-ecological goals, expressed in their own bio-physical terms, across the cleanup and restoration/redevelopment alternatives. Formal social-psychological surveys of potential recreational users and visitors/tourists could measure the relative preferences (importance, acceptance) across the restoration/redevelopment plans (outcomes) under consideration from the perspectives of these important groups. Parallel economic or monetary assessments, perhaps using contingent valuation and or travel cost methods, could extend and cross-validate survey results, and provide dollar-denominated value

indices to facilitate analyses of tradeoffs with development costs and between recreation, tourism and industrial development emphases at a site.

For the Avtex Fibers site, deliberative group processes involving stakeholders and relevant experts (including historians) would have provided an effective approach to identifying the ecosystem and ecosystem service values of most concern to stakeholders. Systematic assessments of ecological values and of historic and sense-of-place values (assuming these were identified as important) are not well developed. Stated-preference monetary assessment methods, such as contingent valuation surveys (methods citation or reference to sec #) might be applied, but there is reason to worry about the sufficiency and the validity of these methods in this context. People have been shown to have considerable difficulty expressing consistent willingness-to-pay estimates for such non-commodity outcomes, and some people find assignments of dollar values to be ethically offensive in this context. Social-psychological "attitude" survey methods could provide reliable and precise relative measures of preferences (importance, acceptance) for any defined population of stakeholders for the array of ecological, historic and sense-of-place outcomes across any defined set of cleanup and restoration/redevelopment options for the site.

Balancing tradeoffs. Because the measures provided by the most common social-psychological survey methods are only relative (across the range of alternatives assessed) this would leave the difficult task of resolving tradeoffs among ecological, historic and sense-of-place values, and between these values and other values and costs, up to the decision maker. Given some consensus that improved biodiversity, habitat quality, energy flow, and/or other biological outcomes were desirable and important to stakeholders, ecological value assessment methods might provide effective and suitable quantitative indices for making comparisons among identified management alternatives in these terms. As with the social-psychological scales, however, ecological assessments would again leave the multi-attribute tradeoff questions to be resolved by the decision maker. This can be an appropriate allocation of decision making responsibilities in many policy contexts, but more sophisticated survey approaches could help to overcome some of the limitations of having only relative measures for multiple value dimensions (attributes), including protection of ecosystems and ecosystem services. Conjoint survey

methods (section Reference) require respondents to explicitly make tradeoffs among multiple value dimensions (attributes), thus revealing the relative contribution of each attribute (in the form of regression coefficients) to relative preferences among the cleanup and restoration/redevelopment plans under consideration.

Managing a site like Avtex Fibers is very complex, with many interrelated and interacting effects for ecosystems and for human society. Thus, a conjoint survey such as that proposed above would most effectively be conducted in the context of an informed, deliberative process, providing a limited set of motivated respondents with expert analyses and information about the inter-relationships among the many potentially competing values at play. For example, respondents would likely require more extensive instruction in the meaning of ecological measures (e.g., biodiversity, energy flow) and how they related to aspects of the actions and outcomes of the alternative management plans than is typically possible in any mass survey approach. In addition, in this context it could be useful for respondents to receive some expert feedback about the possible implications of their expressed preferences for management plans, as effects of environmental changes concatenate through ecosystems and social systems on-site and off, and changing over time. It is important in this context to recognize that the preferences that are derived (constructed) though such an informed deliberative process would not be representative of the reactions of the broader populations of stakeholders. For this reason it may be important to also conduct less intensive survey procedures with larger samples to better predict public response to the plans under consideration and to identify specific public information/education needs that should be addressed in communicating, justifying and implementing the decision.

Communicating outcomes. Additionally, the Agency should advance their capacity to communicate alternative futures and their associated benefits to stakeholders. Representation of scientific information is often obscure to lay audiences. Communicating in terms of the benefits the stakeholders can expect from proposed actions will help focus their interest. Additionally if there are visual ways techniques to represent alternative futures based on different actions it will help stakeholders understand the alternatives from an outcome basis. For example, both monetary/economic and social-psychological assessment methods might make effective

use of perceptual representations (e.g., visualizations of revegetation options as viewed from adjacent homes and prominent tourist and recreation sites and passageways) to improve stakeholders' understanding of the implications of the various restoration/redevelopment alternatives under consideration. In any case the Agency can only benefit from developing communication tools that engage and satisfy the local community's concerns and demonstrates recognition of their preferred outcomes.

Recommendation #5. Develop the capability to utilize valuation techniques to provide a basis to demonstrate performance and communicate the expected or actual outcome from Agency actions.

If valuation concepts and techniques are incorporated early and often throughout the contaminated property redevelopment process then as is suggested in Figure 6: Changing Focus from Remediation to Redevelopment Would Benefit from Increased Integration of Valuation Analysis with Traditional Process Steps, the Agency should be in a position to communicate with interested publics. The expectation is that by effectively integrating consideration of ecological services and their derived benefits into the selection of the remedial and redevelopment actions, managers will be able to communicate "why" they selected the preferred options. Demonstrating to the public that there has been a focus on ecological services that matter to them, and the ability to communicate in terms of the benefits they will derive from the proposed actions, should lead to greater public acceptance of the proposed plan forward.

Additionally, the presence a clearly defined sets of aligned actions and projected benefits should make the selection of performance measures relatively straightforward. Communicating the progress or challenges to such progress as the redevelopment proceeds should be facilitated by having and using performance measures defined in terms of benefits that the interested public understands and accepts as important.

Additionally, the Agency should advance their capacity to communicate alternative futures and their associated benefits. For example, the restoration plan for the Avtex site included replanting and/or encouraging re-growth of three different forest types on appropriate locations within the site. Accurate visualizations of the reforestation projects, including their expected growth over time would be very useful for communicating the implications of alternative plans to stakeholders (only one plan was

actually proposed for the Avtex site), whether in an information context or for systematic value assessments. Achieving and effectively using such visualizations would first require interactions between foresters/forest ecologists and visualization experts (such as some landscape architects) to create accurate and realistic representations of how the different forests would look from significant viewpoints at different stages of the restoration program for each management alternative. Psychologists, communications experts or other relevant social or decision scientists might then be involved in creating appropriate vehicles and contexts for presenting the visualizations to relevant audiences. Technical computer graphics expertise might also be useful in this context. Further interdisciplinary collaboration would be required if the visualizations were to be accompanied by information about expected wildlife or other ecological effects associated with each visualized forest condition. All of this could be a prelude and a perceptual component of a conjoint value assessment survey. The above example may seem a rather intricate process which will require significant time and resources, but keeping in mind that many contaminated properties are under redevelopment for years and in case of Superfund projects decades with proportional resource allocations, this level of effort seems appropriate.

Recommendation #6. Create formal systems and processes to foster an information sharing environment.

Actively document lessons-learned from applications of valuation methods and share broadly among program and project managers. Broad and rapid transfer of experience with integrating valuation concepts and techniques into the process of contaminated site redevelopment should be a lead objective for the Agency. In many ways no two local management situations are exactly alike, so the agency will ultimately build its capacity to utilize valuation to inform its local decisions through a systematic approach of local case-specific demonstrations. The lessons learned from these trial efforts, whether they be successes or failures need to be shared widely across the agency with the regions, program offices and the tool builders in the research organizations. There are a number of ways in which the agency could catalogue and share such experiences, such as reports, databases or BestNets (computer-based networks of users sharing best practices). Obviously the Agency is in the best position to know how to

build off their existing information exchange systems, but however it is done the information should be shared broadly.

Text Box 4: Charles George Landfill

From the late 1950s until 1967, the Charles-George Reclamation Trust Landfill, located 1 mile southwest of Tyngsborough and 4 miles south of Nashua, New Hampshire, was a small municipal dump. A new owner expanded it to its present size of approximately 55 acres and accepted both household and industrial wastes from 1967 to 1976. The facility had a license to accept hazardous waste from 1973 to 1976 and primarily accepted drummed and bulk chemicals containing volatile organic compounds (VOCs) and toxic metal sludges. Records show that over 1,000 pounds of mercury were disposed of and approximately 2,500 cubic yards of chemical wastes were landfilled. The State ordered closure of the site in 1983. That same year, the EPA listed the site on the NPL and the owner filed for bankruptcy. Samples from wells serving nearby Cannongate Condominiums and some nearby private homes revealed VOCs and heavy metals in the groundwater. Approximately 500 people live within a mile of the site in this residential/rural area; 2,100 people live within 3 miles of the site. The nearest residents are located 100 feet away. Benzene, tetrahydrofuran, arsenic, 1,4-dioxane, and 2-butanone, among others, had been detected in the groundwater. Sediments have been shown to contain low levels of benzo(a)pyrene. People face a potential health threat by ingesting contaminated groundwater. Flint Pond Marsh, Flint Pond, Dunstable Brook, and nearby wetlands are threatened by contamination migrating from the site. EPA's involvement at the Site began with groundwater testing conducted by EPA contractor Ecology and Environment, Inc. during 1981 and 1982. The site was proposed for the National Priorities List (NPL) on October 23, 1981, and finalized on the NPL in September 1983. In September 1983 EPA also allocated funds for a removal action at the Site to replace the DEQE's temporary water line with another temporary but insulated water line. Other removal work included construction of a security fence along the northwestern entrance to the landfill, regrading and placement of soil cover over exposed refuse, and installation of twelve gas vents. A remedial investigation and feasibility

study (RI/FS) were also begun in September 1983. The basis for the removal action was documented in the first ROD issued on December 29, 1983.

Text Box 5: Dupage County Landfill

The 40-acre tract of land that is now the Blackwell Landfill was originally purchased by the DuPage County Forest Preserve District (FPD) in 1960 and is centrally located within the approximately 1,200-acre Blackwell Forest Preserve. The landfill was designed to be constructed as a honeycomb of one-acre cells lined with clay. Approximately 2.2 million cubic yards of wastes were deposited in the landfill between 1965 and 1973. The principal contaminants of concern for this site are the volatile organic compounds (VOCs) 1,2-dichloroethene, trichloroethene and tetrachloroethene, detected in onsite groundwater at or slightly above the maximum contaminant level (MCL). Landfill leachate contained all kinds of VOCs and semivolatiles including benzene, ethylbenzene toluene, and dichlorobenzene; and metals such as lead, chromium, manganese, magnesium, and mercury. VOCs and agricultural pesticides have also been detected in private wells, down gradient of the site but at low levels. Some metals (manganese and iron) have been detected above the MCLs in downgradient private wells. Post-remediation, the site now consists mainly of open space, containing woodlands, grasslands, wetlands, and lakes, used by the public for recreational purposes such as hiking, camping, boating, fishing, and horseback riding. There are no residents on the FPD property, and the nearby population is less than 1,000 people. The landfill created Mt. Hoy which is approximately 150 feet above the original ground surface.

Text Box 6: Avtex Fibers Site

The Avtex Superfund site consists of 440 acres located on the bank of the Shenandoah River within the municipal boundaries of Front Royal, VA. The site is bordered on the east by a military prep school (grades 5 -12), on the south by a residential neighborhood, and on the west by the Shenandoah River. From 1940 to closure in 1989 industrial plants on the site manufactured rayon and other synthetics. Tons of manufacturing wastes and by-products accumulated on the site, infiltrated into groundwater under the site and/or

escaped into the Shenandoah River. The Avtex Fibers site was proposed to the National Priorities List on October 15, 1984, and the site was formally added to the list June 10, 1986. EPA began removal activities at the Site in 1989 to address various threats to human health and the environment. The cleanup/restoration plan called for most remaining wastes to be consolidated on site and secured with a protective material (where needed), and a thick soil cover and vegetation (cap).

Front Royal and is located in close proximity to the Appalachian Trail, the Shenandoah National Park and George Washington National Forest, making it a major tourist center for the Blue Ridge Mountains. Biologically, the Avtex site contains some residual forested areas, open meadows and small wetland areas and more than a mile and a half of frontage along the Shenandoah River. The proposed Master Plan for redevelopment, created through a formal Multi Stakeholder Group process, divides the site into three areas: 1) a 240-acre River Conservancy Park along the Shenandoah River combining ecological restoration and conservation of native habitats; 2) a 25-acre Active Recreation Park with boat landings, picnic shelters, and a developed recreational area including a visitor center and soccer fields; and 3) a 165-acre Eco-Business Park, featuring the refurbished historic former Avtex administration building. Clean up of the Axtex site is on-going, and the redevelopment plan is being actively pursued by local government agencies and private industry groups.

Text Box 7: Leviathan Mine Superfund Site

In May of 2000, the EPA added the Leviathan Mine site in California to the National Priority List (NPL) of Superfund sites. The site is currently owned by the State of California, but from 1951 until 1962 the mine was owned and operated by the Anaconda Copper Mining Company (a subsidiary of ARCO) as an open pit sulfur mine. The mine property is 656 acres located in a rural setting near the Nevada border, 24 miles southeast of Lake Tahoe. The physical disturbance from the mine itself is about 253 acres of the property plus an additional 21 acres of National Forest Service land. The site is

surrounded by national forest. In addition, it lies within the aboriginal territory of the Washoe Tribe and is close to several different tribal areas.

The mine has been releasing hazardous substances since the time that open pit mining began in the 1950's. Releases occur through a number of pathways, including surface water runoff, groundwater leaching and overflow of evaporation ponds. In particular, precipitation flowing through the open pit and overburden and waste rock piles creates acid mine drainage (AMD) in the form of sulfuric acid, which leaches heavy metals (such as arsenic, cadmium, copper, nickel and zinc) from the ore. These releases are discharged into nearby Leviathan Creek and Aspen Creek, which flow into the East Fork of the Carson River. Pollution abatement projects have been underway at the site since 1983. Despite these efforts, releases continue today.

The releases of hazardous substances from the mine have significantly impacted the area's ecosystem and the services it provides. In the 1950's structural failures at the mine that released high concentrations of AMD into streams resulted in two large fish kills, and the trout fishery downstream of the mine was decimated during this time. More recently, data have documented elevated concentrations of heavy metals in surface water, sediments, groundwater, aquatic invertebrates, and fish in the ecosystem near the site. This suggests that hazardous substances have been transmitted from abiotic to biotic resources through the food chain, thereby affecting many trophic levels. A recent assessment identifies seven categories of resources potentially impacted by the site: surface water resources, sediments, groundwater resources, aquatic biota, floodplain soils, riparian vegetation, and terrestrial wildlife. The assessment identified five types of ecological services that might be provided by these resources: aquatic biota (including the threatened Lahontan cutthroat trout) and supporting habitat, riparian vegetation, terrestrial wildlife (including the threatened bald eagle), recreational uses (including fishing, hiking, and camping), and tribal uses (including social, cultural, medicinal, recreational and subsistence).

The process of determining compensatory damages and developing a response plan for the site involves a number of different stages for which information about the value of these lost services would be a useful input. For example, in accordance with Natural Resource Damage Assessment (NRDA) regulation under CERCLA (?), the Trustees for the site conducted a pre-assessment screening to determine the damages or injuries that may have occurred at the site and whether a natural resource damage assessment should be undertaken. This requires a preliminary assessment of the likelihood of significant ecological or other impacts from the contamination (corresponding to Step 1 in the process diagram, Figure 2 of this report). The decision was made at that time (July 1998) to move forward with a Type B NRDA, which in principle is a decision to move forward with an assessment of the value of the ecological services that have been lost as a result of the site contamination. A Type B assessment involves three phases: a) injury determination to document whether ecological damages have occurred, b) quantification phase to quantify the injury and reduction in services (corresponding to step 3 of the process diagram), and c) damage determination phase to calculate the monetary compensation that would be required (corresponding to step 4 of Figure 2). In the Leviathan mine case, the Trustees proposed using resource equivalency analysis (REA) based on a replacement cost estimate of the lost years of natural resource services to determine damages for all impacted services other than non-tribal recreational fishing. For this latter ecological service, they proposed using benefit transfer to estimate the value of lost fishing days. Finally, in the decision by EPA about whether to list the site on the NPL and the subsequent Record of Decision (ROD) selecting a final remedy for the site, information about the value of the ecological improvements from cleanup could play an important role, although these decisions are often based primarily on human health considerations.

7. VALUATION IN COLLABORATIVE PARTNERSHIPS

Text Needed

8. ANALYSIS AND REPRESENTATION OF UNCERTAINTIES IN ECOLOGICAL VALUATION AND COMMUNICATION OF ECOLOGICAL VALUATION INFORMATION

8.1. Uncertainty

8.1.1 Introduction

Ecosystem valuation efforts are inevitably subject to uncertainty, which is by no means limited to a single method or group of methods of ecological valuation. Rather, uncertainty is the rule, not the exception in this domain. Assessments of uncertainty allow more informed evaluations of proposed policies and comparisons among alternative policy instruments, and unless uncertainty is taken into account and thoughtfully conveyed to decision makers, the ultimate usefulness of respective assessments may be compromised. Because any given policy may result in a range of different outcomes, decision makers must be provided with sufficient information about the distribution of probabilities so that they can take uncertainty into account in their policy choices. Whether decision makers wish to adhere to maximizing expected utility, avoiding major risks through a "maxi-min strategy," or some other decision principle, they have to incorporate the uncertainty that policy choices always entail. In addition, if the sources of key uncertainties are not identified, an opportunity is lost to develop potentially important insights regarding the design of research strategies to reduce uncertainty in future analyses.

It is important to distinguish between precision and accuracy, and likewise between imprecision (uncertainty) and inaccuracy. In a quantitative, statistical context, the contrast is captured, respectively, by measures of variance and bias. Our focus is exclusively on uncertainty, that is, precision or the lack thereof. Some existing and potential methods of ecological valuation may involve systematic bias, but that is not the subject of this essay. An exception is that we note that individual methods of ecological valuation may display inherent tradeoffs between precision and accuracy. For example, an attempt to make questions (and the implied budget constraint) in a contingent valuation analysis more realistic and credible can have the desirable effect of increasing precision (reducing uncertainty and variance in responses) while also having the

unfortunate effect of bringing about truthful revelation of preference problems, that is, introducing strategic bias.

Reflecting on the role of uncertainty in ecological valuation, three key questions arise. First, what are the major sources of uncertainty? More specifically, what types of uncertainty are likely to be most important with alternative valuation methods for specific applications? Second, what methods are available to characterize and communicate uncertainty in the results of ecological valuations? Here we are interested not only in the formats that can be employed – such as confidence intervals, probability distributions, and pictorial representations – but also the types of interactions between analysts and policymakers that can be employed to convey uncertainty most effectively. Related to this, how can uncertainty in ecological valuations be incorporated in various decision-analytic frameworks, such as benefit-cost analysis, cost effectiveness analysis, and multi-criteria analysis? This raises the question of whether specific institutional obstacles to conveying uncertainty currently exist (and, if so, how they can be overcome), and likewise whether some institutional attributes currently favor the analysis and communication of uncertainty in ecological valuations. A third and final key question is associated with the types of research - data collection, improvements in measurement, theory building, theory validation, and others - that can be pursued to reduce uncertainty for particular sources in specific applications.

Section 6.1.2 describes the major sources of uncertainty in ecosystem and ecosystem services valuation. Section 6.1.3 examines the potential for uncertainty assessment of ecological values, describing both the merits of formal quantitative uncertainty assessments and the additional efforts that would be required for government agencies to carry out such assessments. Section 6.1.3.4 focuses on issues associated with communicating uncertainty to policy makers (and other audiences) in ecological valuations. Section 6.1.5 assesses the potential value of uncertainty assessments to the research agenda of the U.S. Environmental Protection Agency and other researchers.

8.1.2 Sources of Uncertainty in Ecological Valuations

Valuation of the benefits of proposed public policies entails three analytic tasks, each potentially subject to uncertainty: forecasting biophysical outcomes, forecasting socio-economic reactions to these outcomes, and valuing the consequences of all of these

changes. In addition, the application of valuation methods per se also introduces uncertainty, in terms of the appropriateness of the assumptions, the comprehensiveness of information captured by the method, the representativeness of the sample, etc. It might be tempting to limit attention to the uncertainty of valuation per se, but all of these sources of uncertainty are of potential importance, and there is no reason – on the basis of theory alone – to judge one more important than the other a priori. Rather, the relative magnitude of the uncertainty involved in these essential steps in the valuation process is fundamentally an empirical question.

Uncertainty of biophysical changes and their impacts. The first stage of valuation for policy analysis – predicting the biophysical impacts of some public policy (relative to a predicted business-as-usual biophysical baseline) — typically involves three conceptually distinct but interrelated sources of uncertainty:

a) limitations due to lack of information and data quality (lack of data, faulty data, or data of variable quality in particular contexts);
b) stochastic (random) variation, also known as within-model uncertainty (i.e., variation beyond analysis, such as temperature fluctuations caused by solar flares); and
c) theory limitations (across model uncertainty).

At the bio-physical level any characterization of current (or past) ecological conditions will have numerous interrelated uncertainties, and these uncertainties will be magnified and added to by any effort to project future conditions, with or without some postulated management action. Ecosystems are complex, dynamic over space and time, subject to the effects of stochastic events (such as weather disturbances, drought, insect outbreaks, fires, etc) and our knowledge of these systems is incomplete and uncertain. Errors in projections of future states of ecosystems are thus unavoidable, and constitute a significant and fundamental source of uncertainty in any assessment of ecosystems/services benefits.

While the currently available methods for dealing with uncertainty may be sufficient for some simple evaluation problems, the valuation of changes in ecosystems and ecosystem services raises issues not well addressed by any existing methods. For

example, at the biophysical level it is extremely difficult or completely unclear how to calculate the uncertainty (error) in the projection of even a single outcome (endpoint) from a complex ecological system composed of multiple interacting variables that may be separately non-linear and collectively subject to the influence of external stochastic events. Modeling methods, such as sensitivity analyses, may be used to estimate the range of possible outcomes (or at least best-case, worst case extremes) for a single endpoint, but even this approach becomes unwieldy when the outcomes relevant to the value assessment are themselves composed of multiple interrelated variables.

Uncertainty of socio-economic reactions and their impacts. In addition to the uncertainty introduced by biophysical projections, uncertainties in the individual and social consequences of ecological changes arise from the fact that we are rarely able to determine with certainty the individual or social consequences of a given change in ecosystems/services. Thus the second stage of valuation – predicting the socio-economic reactions to biophysical impacts and the consequences of these reactions – is subject to the same three sources of uncertainty. Regarding information and data quality every measure of economic activity, every inventory of crops and herds, and every survey is subject to some degree of error. Regarding random variation, socio-economic reactions are sensitive to the confluence of unrelated events whose co-incidence is intrinsically unpredictable. For example, a key vote on an important piece of environmental legislation may be scheduled for a date that turns out to follow a highly-publicized oil spill. Or a community's decision to relocate housing away from an increasingly flood-threatened area may be derailed by a national economic downturn that reduces matching funds for the relocation program. In addition, significant policy decisions are often made by single individuals or very small sets of individuals, whose personal idiosyncrasies create uncertainty in prediction quite parallel to the solar flares. Regarding theory limitations, every social, economic or political forecast is based on implicit or explicit theory of how the world works, represented either by the "mental models" in the minds of the forecasters or in the formal and explicit methods used in econometric modeling, systems dynamics modeling, etc. Theories and their expressions as models are unavoidably incomplete, and of course may simply be incorrect in their assumptions and specifications.

uncertainty arising from the application of valuation methods. Valuation methods per se are also subject to data and theory limitations. They unavoidably rely on assumptions that introduce uncertainty. Revealed-preference approaches, which seem more reliable than stated-preference approaches because of the hypothetical nature of the questions that the latter employ, are prone to some degree of model mis-specification and incompleteness. Consider the hedonic pricing method and its typical reliance on regression analysis that requires assumptions of either linear or specified curvilinear relationships. The choice of the assumption may be the best among the alternatives, but whichever assumption is chosen will not fit the data with total accuracy. The price data for the analysis is also subject to some degree of measurement error, and, insofar as hedonic pricing only reveals choices to secure private utility, its capture of existence values and the utility that comes from "public regard" is incomplete. The results of stated-preference approaches, in addition to the problem of the hypothetical nature of the questions, are also subject to uncertainty as to the impact of variations in how questions are posed, how they are interpreted by different respondents, and how reactions are influenced by differences in the information that is presented.

In addition, all assessments of expected consequences are about anticipated, not experienced satisfaction those consequences might bring. To take a simple example, the choice of a vanilla ice cream cone over chocolate is based on the anticipation that consuming the vanilla will bring greater pleasure/satisfaction than the chocolate (and perhaps even further that a pleasant gustatory experience will contribute toward a more ultimate goal of improved well-being, happiness in life or self actualization). In fact research has shown that even in relatively simple and familiar situations people err considerably in their anticipation of the satisfaction they will attain from a given outcome. When the values and choices at issue are about imperfectly projected changes in ecosystems/services, where previous experience is limited and where the time horizons are much greater, there is even less certainty in the accuracy of anticipated satisfaction. These anticipation errors become even more problematic in the typical circumstances of an environmental management decision, where the goals and the intended beneficiaries are some loosely defined society, some members of which may not yet exist, and only a small number of whom are involved in any direct way in the consideration and decision

making process. In such contexts any notion of a final and accurate assessment of the true value of some change in ecosystems/services must be illusory. Still, people and agencies must continue to evaluate alternatives and make decisions based on their best estimate of what consequences will follow and how they will contribute to proximate and ultimate goals.

Uncertainty in benefits transfer. In addition, even if existing estimates are developed using an appropriate model, analysts are often required to apply them to contexts that differ from those in which they were developed. The possibility that appropriate adjustments are not made in transferring estimates to different contexts introduces another source of uncertainty. In order to identify the types of uncertainty most likely to be at issue for individual valuation approaches in specific contexts, two issues are relevant: the sensitivity of an approach to the potential sources of uncertainty listed above, and the magnitude of uncertainty thereby generated. The consequence of data limitations can be assessed by sensitivity analysis to determine the variation in results implied by variations in data. Vulnerability to theoretical limitations is more difficult to assess, but can be gauged - in some cases - by sensitivity analysis with alternative models. The consequences of stochastic variation can be assessed in simple models by observations of measures of dispersion (for example, variance of estimates), whereas in more complex models, stochastic simulations - Monte Carlo analysis - can be employed (Jaffe and Stavins 2004).

8.1.3 Approaches to assessing uncertainty.

The tasks of assessing the uncertainty of the elements that go into a valuation involve estimating a distribution (rather than a single point) of values arising from the combined uncertainties of the elements of the analysis, and a diagnosis of the elements that are contributing most heavily to spreading this distribution. Given the multiple levels of elements that can add to uncertainty, the most complete approaches will be unavoidably complex themselves.

Monte Carlo analysis as an approach to the formal uncertainty assessment of ecological values. Due to the number of sources of uncertainty in many ecological valuations and the complexity of their interactions, assessments of the extent of uncertainty that are conducted without formal quantitative analyses are unlikely to

represent accurately the true extent of uncertainty. No sensitivity analysis or expert judgment is likely to be able to account for the implications of all the sources of uncertainty in inputs, which can be incorporated in a Monte Carlo analysis. Therefore over the years, the use of formal quantitative uncertainty assessment, and in particular Monte Carlo analysis, have been shown to provide reliable and rich characterizations of the implications of uncertainty, and therefore have become common in a variety of fields, including engineering, finance, and a number of scientific disciplines.

Monte Carlo analysis has also been found to be useful in certain policy contexts. In particular, the U.S. Environmental Protection Agency (EPA) recognized as early as 1997 that it can be an important element of risk assessments (U.S. Environmental Protection Agency 1997). But efforts to formally quantify uncertainties rarely have been made in the context of ecological valuations. More often, uncertainty has been addressed qualitatively or through sensitivity analysis.

Monte Carlo analysis can be implemented with relative ease. The first step is the development of probability distributions of uncertain inputs (to an ecological valuation or other analysis), where the probability distributions reflect the implications of uncertainty regarding respective inputs for the range of its possible values and the likelihood that each value is the true value. Once probability distributions of inputs to an ecological valuation are established, a Monte Carlo analysis determines the resulting probability distribution of the valuation by carrying out the calculation thousands, or even millions, of times. With each iteration of the calculations, new values are randomly drawn from each input's probability distribution and used in the calculation of the ultimate ecological valuation. Over the course of these iterations, the frequency with which any given value is drawn for a particular input is governed by that input's probability distribution. If a sufficient number of iterations are performed, the range of resulting valuation estimates and the frequency of particular estimates within that range can be used to determine the probability distribution of values arising from those input uncertainties that have been characterized in the analysis.

As it is unlikely that a Monte Carlo analysis will comprehensively address all sources of uncertainty in the estimation of ecological values, even the results of such an analysis will likely understate the range of possible outcomes that could result from a

related public policy. Yet the ranges produced by such an analysis would still provide more reliable information about the implications of known uncertainties. In turn, these ranges can better inform judgments by policymakers as to the overall implications of uncertainty for their decisions.

Monte Carlo analysis also provides information on the likelihood of particular outcomes within a range. Indeed, an understanding of the likelihood of values within a range is essential to any meaningful interpretation of that range. Without such an understanding, inappropriate conclusions may be drawn from the presentation of a range of possible outcomes. For example, when a range of possible ecological values is provided, some may assume that all values within that range are equally likely to be the ultimate outcome. But this is rarely the case. Others may assume that the distribution of possible values is symmetric. This, too, often may not be the case.

Typical methods of addressing uncertainty, such as less systematic sensitivity analysis, often cannot provide meaningful guidance as to the likelihood that ecological values will exceed or fall below certain values. In fact, such analyses can sometimes inadvertently provide misleading suggestions as to the likelihood of certain outcomes. These analyses can indicate the extent to which uncertainty in particular inputs contributes to overall ecological values, but the implications of uncertainty in one input cannot be put into context without the use of a formal quantitative assessment of uncertainty to characterize that overall uncertainty. Absent such an assessment, there is a risk that results may be perceived incorrectly as providing information regarding overall uncertainty in ecological values.

In some circumstances uncertainty assessments can change point estimates. Monte Carlo analysis can reveal when uncertainties in inputs to an ecological valuation cause the expected value to differ from what would be suggested by a deterministic analysis. Less systematic approaches may explore the implications of uncertainties that can bring about such results, but cannot address all possible outcomes resulting from those uncertainties or indicate the probability associated with any one outcome.

Implementation of a Monte Carlo analysis imposes two requirements that are not strictly necessary to develop point estimates. First, instead of requiring a single point estimate for each input, Monte Carlo analysis requires the development of probability

distributions for important, uncertain inputs. Second, numerous repetitions of the calculations of the ecological value must be performed. These requirements may appear burdensome, but to a large degree, they can entail little additional effort, relative to what is already expended in many ecological valuations. Furthermore, as with any ecological valuation, a Monte Carlo analysis does not need to be exhaustive to offer valuable insights.

In developing probability distributions for uncertain inputs, uncertainty from statistical variation can often be characterized with little additional effort relative to that needed to develop point estimates. Much of the data necessary for such characterizations already will have been collected for the development of point estimates. Characterizing other sources of uncertainty in inputs can require more effort. For example, expert elicitation methods can be employed. These methods are formal, highly structured, and well documented procedures whereby expert judgments are obtained (Morgan and Henrion 1990; Cleaves 1994).

The amount of additional effort necessary to develop a Monte Carlo analysis can be minimized through careful consideration of which input uncertainties are worthwhile addressing in the analysis, since valuable insights can be gained even if the uncertainties in only a few inputs are characterized. Evaluation of how an input factors into an analysis and a preliminary assessment of uncertainty may make clear that efforts to characterize uncertainty in the input would have little affect on the findings of a Monte Carlo analysis. Thus, significant efforts to characterize uncertainty in that input would not be warranted. Such an assessment also could lead to the opposite conclusion, thereby justifying additional effort. And some inputs may be significant elements of numerous ecological valuations, providing additional justification for efforts to develop more complete characterizations of uncertainty in their values.

While a Monte Carlo analysis can require additional effort to characterize uncertainty in inputs to an ecological valuation, that effort often may be warranted even in the absence of the needs of a Monte Carlo analysis. This is because such characterizations of uncertainty may be necessary just to develop an accurate point estimate for an input. If a point estimate represents an input's expected value, the development of that point estimate requires an implicit judgment about that input's

probability distribution. Characterizations of uncertainty required in a Monte Carlo analysis simply make those implicit judgments explicit. Therefore, in addition to making possible quantification of uncertainty in the results of an ecological valuation, these characterizations can improve the empirical basis for, and quality of point estimates used as inputs to the analysis.

Developments in computer performance and software over the years have substantially reduced the amount of effort required to conduct calculations for a Monte Carlo analysis, once input uncertainties have been characterized. Widely available software allows the execution of Monte Carlo analysis in common spreadsheet programs on a desktop computer. Also, modern programming techniques allow the writing of Monte Carlo computer programs with minimal additional effort, relative to that needed to produce point estimates.

Expert elicitation for gauging and conveying uncertainty. A host of "expert elicitation" methods, described in Section Y.Z of this report, can provide indications of uncertainty as well as estimates and forecasts by the experts involved. In its very simplest form, a single expert's assessment of the uncertainty of his or her estimate, forecast, or valuation can be provided, whether it is based on implicit judgment or a more explicit approach like the Monte Carlo technique. Policymakers can elicit more information from the expert, such as the assumptions underlying his or her analysis or the bases for uncertainty, in order to get a deeper understanding of the reliability of the expert's input and the nature of the uncertainty. However, the bulk of expert elicitation methods involve multiple experts, who may or may not be brought into interaction with one another. Because eliciting the input from multiple experts permits compiling and comparing their judgments, expert elicitation can be used to assess the disagreement among experts. If the experts are of equal credibility, such that none of the judgments can be discarded in favor of others, the range of disagreement reflects uncertainty. That is, if top scientists express strong divergences in their estimates, forecasts, or valuations, the existence of a high level of uncertainty is irrefutable. However, this is an asymmetrical relationship, in that narrow disagreement does not necessarily reflect justified certainty—the experts may all be wrong in the same direction, which is not uncommon in light of the fact that experts are often paying attention to the same

information and operate within the same paradigm for any given issue (Ascher & Overholt, 1984: 86-87). When experts are brought into some form of interaction prior to providing their final conclusions (e.g., by exchanging estimates and adapting them in reaction to what they learn from one another), the errors due to incompleteness (e.g., biologists may be unaware of atmospheric trends that information from atmospheric chemists could redress) can be reduced. However, such interactions run the risk of "groupthink" – unjustified convergence of estimates due to psychological or social pressures to come closer to agreement (Janis, 1982).

For many expert elicitation methods, translation into probabilities is difficult. For example, simple compilations of estimates (e.g., contemporaneous estimates of species populations) from different experts will provide a table with the range of estimates, but will not convey the degree of uncertainty that each expert would attribute to his or her estimate, and the compilation in itself cannot generate this information. In contrast, a compilation of estimates that come with confidence intervals could provide this information. To take another example, the well-known Delphi technique (described in some detail in Section Y.Z) typically provides the inter-quartile range (i.e., the middle 50% of estimates) following several rounds of exchanges of estimates, but does not provide probabilities reflecting the uncertainties of each estimate or of the set of estimates.

8.1.4 Communicating Uncertainty in Ecological Valuations

In order to assess how much confidence to attribute to the projections involved in the valuation, decision makers must also be informed about the analyst's own judgment of the uncertainty of the valuation and its prior steps, and the assumptions underlying the valuation analysis. Making decision makers aware of these assumptions is also important because decision makers often have to explain and justify their decisions by clarifying the assumptions driving the analysis.

In order to convey to policy makers the degree of uncertainty in an ecological valuation, the simplest expressions - whether quantitative (measures of dispersion, such as variance) or qualitative (such terms as "likely," "very likely," etc.) - are typically inadequate. Analysts can specify the central tendency of an estimate (mean or median value, as appropriate) plus a confidence interval (for example, the 95% confidence

interval), but in some cases this may require possibly arbitrary judgments on the part of the analyst (Moss & Schneider 2000). Furthermore, providing policy makers with such ranges of results can be highly misleading, because those without training in probability and statistics may be likely to assume - in effect - that the probability distribution of values between the end-points is uniform, which is rarely, if ever, the case. Sensitivity analysis can help in this regard, although what is really needed is a description - verbal or pictorial - of the full probability distribution.

Institutional obstacles to conveying uncertainty may be related to the understandable reluctance of analysts to expose themselves and their work to the risk of appearing to be lacking in rigor. Analysts may thus have an unfortunate incentive to exclude or otherwise downplay components of their analyses that they fear may jeopardize the credibility of their overall effort. Suppressing less certain information runs counter to the need for transparency and the reality that all estimates have some degree of uncertainty (Arrow et al. 1996).

Historically, efforts to address uncertainty in ecological valuations - and more broadly, in Regulatory Impact Analyses (RIAs) - have been limited. But guidance set forth in the U.S. Office of Management and Budget's (OMB) Circular A-4 on Regulatory Analysis in 2003 has the potential to enhance the information provided in RIAs regarding uncertainty.

In the past, point estimates have been given far greater prominence in RIAs and other government valuations than discussions of uncertainty associated with them. Uncertainty assessments are often relegated to appendices and discussed in a manner that makes it difficult for readers to discern their significance. This is perhaps inevitable given that single point estimates can be communicated more easily than lengthy qualitative assessments of uncertainty or a series of sensitivity analyses. The ability of Monte Carlo analysis to produce quantitative probability distributions provides a means of summarizing uncertainty that can be communicated nearly as concisely as point estimates. The need for and means of communicating uncertainty in such a fashion has been addressed in the existing literature. If a summary of uncertainty in an estimate is not given prominence relative to the estimate itself, context for interpreting the estimate and opportunities to learn from uncertainty associated with it may be lost.

Some resistance to the use of formal uncertainty assessments such as through Monte Carlo analysis and prominent presentation of the results may be due to the perception that such analysis requires more expert judgment and therefore makes the results presented more speculative. Also, some might argue that, given the inevitably incomplete nature of any uncertainty analysis, prominently presenting its results would incorrectly lead readers to conclude that results of an ecological valuation are more certain than they are. Both concerns seem to be unfounded. First, as described above, developing characterizations of uncertainty (such as for inputs in a Monte Carlo analysis) often simply involves making explicit and transparent expert judgments that necessarily already must be made to develop point estimates for those inputs. Moreover, to the extent that an uncertainty analysis is thought to be incomplete in its characterization of uncertainty, that fact can surely be communicated qualitatively.

8.1.5 Decision-making with uncertainty

Understanding how information about values will and should be used by decision-makers is crucial for understanding how the valuation analysis should be conducted and how uncertainty should be conveyed.

Static decision-making with uncertainty. When there is incomplete information about how policy will affect ecosystems and the provision of services or how such changes affect human well-being, decision-makers will face making choices under uncertainty. This section outlines several methods for incorporating uncertainty into decision-making.

One approach for decision-making under uncertainty is to assign probabilities for all possible outcomes and then to make decisions based on expected net benefits. For example, suppose the benefits of reducing air emissions depend on whether it will be sunny or cloudy. The probabilities of each event (sunny, cloudy) will be multiplied by the net benefits in each case to derive the expected net benefits. When there are both present and future benefits and costs the approach would be to calculate expected net present value. Maximizing expected (present value) net benefits has the advantage of being transparent and straight-forward. The disadvantages of this approach are that it requires probability assessments and assumes risk neutral behavior.

When faced with uncertain outcomes, many people would rather choose an option with a lower risk of bad outcomes even though the expected net benefits of this option are the same or lower than some other option. For example, in choosing between an option that gives $100 for sure versus a 75% chance of no return and a 25% change of $400, the vast majority of people choose the $100 option (Kahneman and Tversky 2000). Both options have an expected return of $100 but they differ in the amount of risk that a person faces. One could ask whether people would rather receive $90 or $80 versus the 75% chance of no return and a 25% change of $400. The difference between the sure thing and the expected value of the risky option that makes people indifferent between the two choices is called the risk premium. People willing-to-pay a risk premium are said to be risk averse. It may the case that people are risk loving rather than risk averse, as when they buy lottery tickets. When people are asked whether they would rather choose an option in which they lose $300 for sure or an option in which there is a 25% chance of losing nothing and a 75% chance of losing $400, the majority of people choose the risky option (Kahneman and Tversky 2000). Note that this is really the same gamble as the first one where this time it is framed in terms of losses rather than gains, showing the importance of framing of risk on decision-making.

Incorporating risk aversion (or risk loving) behavior into decision-making requires changing the objective from one of maximizing expected net benefits to maximizing expected utility, where the latter incorporates attitudes toward risk. In practice, it may be difficult to get an accurate assessment of attitudes toward risk making it difficult to make expected utility operational.

In many cases, it is not possible to establish objective probabilities based on prior experience or first-principles. For example, novel situations (designing a high-level nuclear waste repository) or situations in which there is a regime shift in system structure mean that past experience cannot be relied upon to generate objective probabilities. In such cases, probability assessments must be subjective. For individual decision-making, subjective probabilities are whatever the individual assesses the odds to be. For policy purposes, subjective probabilities will need to be set by some means, either through asking experts or surveying the affected public.

In instances of truly novel events, people may be unable or unwilling to assign even subjective probabilities. It may also be difficult to even know what outcomes might be possible. Some methods for decision-making under uncertainty do not rely on probability assessments. Rather than trying to maximize expected net benefits or expected utility, methods such as safe minimum standards or the maxi-min rule focus on minimizing the risk of a very bad outcome occurring. Under safe minimum standards decision-makers should avoid decisions that might push a system beyond a threshold that could lead to large negative consequences, unless the costs of doing so are intolerable. Under maxi-min rules, society should choose among alternatives based on the alternative that generates the best worst-case outcome. Maxi-min rules can be justified either on the basis of extreme risk-aversion or in cases where it is not possible to assign probabilities to outcomes. Maxi-min rules have been justified as being an appropriate strategy in cases where probabilities for events cannot be assigned (Arrow and Hurwicz 1972, Maskin 1979).

A related notion in environmental policy circles is the Precautionary Principle. The Precautionary Principle states that society should avoid actions that may result in large damages even though there is not conclusive scientific proof of cause-effect relationships. Precautionary Principle language is included in many international treaties including Agenda 21 from the 1992 United Nations Conference on Environment and Development.

<u>Dynamic decision-making under uncertainty</u>. Most policy decisions that affect ecosystems and services have potentially long-lasting effects, can be revisited in the future, and are subject to considerable uncertainty. For these reasons it is important that the value of protecting ecosystems and service incorporate both dynamics and uncertainty. However, doing so makes the valuation exercise considerably more difficult. It is not sufficient to do analysis about the current situation and do valuation studies about the current situation. Rather, what is needed is information about the likely state of the ecosystem and the provision of services through time and consequent associated values through time.

An approach to dynamic decision-making is to choose a series of decisions that maximize the present value of expected net benefits or expected utility. In principle,

optimal decisions can be derived by applying optimal control theory or dynamic programming. When there is little opportunity or it quite costly to revisit decisions, the path of decisions may be chosen at the outset, what is called an open-loop strategy. Without uncertainty, open-loop strategies can be optimal. However, with uncertainty and learning through time, new information may reveal that a change in plan is needed. Closed-loop (feedback) strategies allow decisions through time based on the information available at that time. Closed-loop strategies can be ex-ante optimal, i.e., optimal given the information at the time of the decision, even with uncertainty. Stochastic dynamic programming is the mathematical tool that can be applied to finding optimal closed loop strategies. A drawback to the practical application of stochastic dynamic programming is that the complexity of the analysis rises exponentially with time and potential alternative choices considered (the curse of dimensionality). This difficulty may require that analysts adopt rules of thumb or considerably simplify the problem in order to make headway.

An approach that tries to be both practical and to incorporate the ideas of stochastic dynamic programming is adaptive management. In adaptive management, current management actions are designed partly as experiments to reduce uncertainty and provide a broader base of knowledge that contribute to more effective future management decisions (Walters 1986). There are tradeoffs in experimentation: expected benefits in the near term may have to be sacrificed in order to learn information that may be of value for future decisions. There may also be institutional impediments to using management decisions as experiments or in changing management decisions too often.

Special considerations arise when outcomes are irreversible (e.g., species extinctions), or reversible only at great cost (e.g., ecosystem restoration), and when there is uncertainty, particularly about how the future generations might value benefits. In such cases there is value to preserving flexibility and avoiding irreversible or difficult to reverse decisions until uncertainty is resolved. The value of avoiding irreversible outcomes is called option value, or in some literature quasi-option value (Arrow and Fisher 1974, Henry 1974). The importance of avoiding irreversible outcomes (or accounts that can be reversed only at some cost) when the passage of time reduces uncertainty can be illustrated with a simple example based on Arrow and Fisher (1974). They considered a two-period model in which there is a choice between developing and

preserving land. If land is preserved in the first period, then there will again be a choice between developing and preserving land in the second period. Development, however, is assumed to be irreversible so that if land is developed in the first period there is no choice in the second period. Suppose that the benefits of development are 100 per period. Suppose that the current benefits of preservation are 90, while the second period benefits of preservation are uncertain with a 50% chance of being 160 and a 50% chance of being 20. For simplicity assume there is no discounting of future benefits. If one were making a one-time decision with a goal of maximizing expected present net benefits, then the choice would be to develop in the first period rather than preserve. Developing in the first period would give expected benefits of 100 + 100 = 200, versus expected benefits of preservation of 90 + (0.5x160 + 0.5x20) = 180. However, choosing to preserve in the first period allows the flexibility to making a choice over whether to preserve or develop after it is learned whether preservation is highly valued or not. If preservation value turns out to be high in the second period then the land can continue to be preserved. If preservation value turns out to be low in the second period then the land can be developed. In this case, the expected net benefits by choosing to preserve in the first period are 90 + (0.5x160 + 0.5x100) = 230. Taking account of the value of preserving options means that it is optimal to make the choice to preserve in the first period. This type of theory has been well developed in financial theory (see for example Dixit and Pindyck 1994).

Other notions of how to manage complex ecosystems over time focus on system properties such as stability or resilience of the system rather than attempting to optimize present value of expected utility derived from the system. Stability may be desirable because it is costly to adjust to variable flows of ecosystem services. Lack of stability might also cause fundamental shifts in ecosystem state to less desirable conditions (Carpenter et al. 1999). This latter notion is related to system resilience. Resilience can be defined as a measure of the ability of a self-organized system to absorb shocks and disturbances and remain in a desirable state. Management actions should be designed to increase system resilience, both biophysical and social, and build capacity for learning and adaptation. A second definition of resilience is the speed with which a system returns toward an equilibrium state.

8.1.6 Contributions of uncertainty assessment in guiding research initiatives

Assessments of the magnitude and sources of uncertainty can help establish research priorities and whether policy changes should be delayed until research reduces the degree of uncertainty associated with possible changes. Determining whether the major source of uncertainty is in weak data, weak theory, randomness, or inadequate methods can help to guide the decision on how to allocate scarce resources for research, or whether further research is worth pursuing. Even stochastic uncertainty can sometimes be addressed by initiating research that focuses on factors previously treated as exogenous to the theories and models. For example, an earthquake-risk model based on historical frequency will have considerable random variation due to the exclusion of detailed analysis of fault-line dynamics; bringing fault-line behavior into the analysis may lead to reductions in such uncertainty (Budnitz et al. 1997).

Using uncertainty analysis to guide research priorities requires, of course, sensitivity to the feasibility of filling the gaps. Some data needs are simply too expensive to fulfill, and some methods have intrinsic limitations that no amount of refinement will fully overcome.

Uncertainty assessment can also provide insight into whether near-term progress in reducing uncertainty is likely, based on its sources and the feasibility of addressing these limitations promptly. However, it is important to avoid the pitfall of delaying actions to address problems simply because some uncertainty remains – it always will.

8.2. Communication and valuation

Three essential functions of communication in the context of valuing the protection of ecological systems and services are:

a) communication within the valuation process itself;
b) communication of resulting values to inform decision-making; and
c) communication of the results of the valuation and decision-making processes to stakeholders and others.

Understanding how information about values will and should be used by decision-makers is crucial for understanding how the valuation analysis should be conducted and its results conveyed, including how uncertainty should be conveyed.

Within the valuation process itself, how decision objectives, decision attributes, and specific measures of values are communicated can determine the outcome of the process. Good communication practices include the use of an analytic-deliberative process, in which analysis and deliberation occur iteratively and interactively (NRC, 1996). The valuation process (see Figure 2) includes iterative problem definition and description by stakeholders, to clarify what and whose values will be represented in the valuation process. Communication of resulting values to inform decision making is simplified to the extent that decision makers or their representatives are involved in the process. Participative, analytic-deliberative processes are discussed at length in Chapter 4.

Values, decision objectives, and decision attributes can each be defined either qualitatively or quantitatively, and represented in a wide variety of ways. Several critical design choices are likely to influence the effects of communicating values to parties not involved in the valuation process, either to inform decision-making or to share results with other interested parties. In communicating VPESS, key choices include (1) how to describe the ecological functions, systems and services to which the valuation pertains, (2) how to express values most meaningfully - whether to quantify or use non-numerical representation of values, use of visual and narrative strategies for each, and interactivity and related choices regarding the medium and mode of information presentation, and (4) uncertainty. Those choices will in turn either facilitate or hinder specific kinds of deliberations and analyses. Finally, evaluating communications is critical to understanding their effects and improving them.

Decision making in public policy often requires translation and/or aggregation, from one specific context to another, or from one level of decision making (e.g., local) to another (e.g. regional), and inevitably involves value trade-offs. Specific choices of how to represent or communicate values will influence the ease and transparency with values can be translated or aggregated, and with which trade-offs can be made. As discussed in Chapter 4, values that are quantified (e.g., monetized) may be easier to aggregate or

compare than those represented qualitatively; use of multiple metrics is likely to complicate aggregation and comparison.

8.2.1 Describing ecological functions, systems and services

The focus of the value discussion in the NRC report (2001) and SAB review of the ROE (US EPA SAB 2003) and related literatures (e.g., Failing and Gregory, 2003) is not on dollars per se, but on ends and decision or management objectives - that is, qualitative expressions, and a wider variety of expressions of value - not just monetary expressions of value.

Communicating the value of protecting ecological systems and services requires conveying not only value information, but also information about the nature and state of the ecological systems and services to which they apply. The latter can be and is often conveyed using mapped ecological information, other visualizations including photographs and graphs, ecological indicators, and narratives. Integrated models with a geospatial interface are another approach to depicting the state of ecological systems and services. The US EPA SAB has proposed this kind of framework for reporting on the condition of ecological resources (U.S, EPA Science Advisory Board, 2002). The Report on the Environment (ROE) and REMAP reports illustrate a range of representational approaches.

It's critical to communicate ecological processes as well as static information or states; the EPA Science Advisory Board review of the ROE (US EPA SAB 2005) and several other reports (e.g., Schiller et al., 2001; Carpenter et al., 1999; Janssen and Carpenter, 1999) make the point that people need to understand the underlying causal processes, to understand how ecological changes affect things they value (e.g., ecological services).

Related issues of scale and aggregation are also important. Both the NRC report (2001) and the SAB review of the ROE (2005) emphasize the importance of using regional and local indicators - of not aggregating information data to the point where it obscures critical ecological threats or problems. The SAB states that "some environmental changes are best understood by considering regional impacts" (EPA SAB, 2003). Further, while some authors recommend simple summary indicators (e.g., Schiller

et al., 2001; Failing and Gregory, 2003); others emphasize disaggregating indicators (US EPA SAB 2003).

Reporting on the nature, state of and changes in ecological systems and services is a key component of value elicitation and communication, but needs to be married with equal consideration of how to convey the value of protecting them.

8.2.2 Communicating values meaningfully

Communicating values is complicated by the likely disconnect between popular and other uses of the term "values" and what economists in particular mean by the term. Variously described by sociologists and social psychologists as beliefs, goals, or even cultural imperatives, stable sets of values (e.g., benevolence, self-direction, security, hedonism and others) have been identified across cultures, although values vary with history and culture (Hitlin and Piliavin, 2004). Conservation versus openness to change, and self-enhancement versus self-transcendence are two dimensions identified as underlying values (Schwartz, 1994). Values are sometimes conflated with attitudes (which are positive or negative evaluations of an object), traits (which are enduring attributes of personality), norms (which are situation specific) or needs (which are biophysical influences on behavior) (Hitlin and Piliavin, 2004). All of these concepts differ from the sense in which economists use the term value, although attitude appears to come closest.

Value elicitation includes contingent valuation and attitude judgments (generally on rating scales, but also using ranking tasks). It also includes qualitative expressions and narrative expressions of value, defined by the identification of associated ends, and the means to achieve those ends.

Context and framing influence how people rank, rate and estimate values (Hitlin and Piliavin, 2004; Horowitz and McConnell, 2002), as well as the interpretation of all kinds of value-related information (add Slovic and other references). Decision makers and others come with their own prior beliefs and attitudes, of which communicators should be aware (Morgan et al, 2002). To support decisions effectively, it is critical that communications be designed to address the recipient's goals and prior beliefs, taking into account likely context and presentation effects. For example, linear graphs are likely to convey trends more effectively than tables of numbers (Shah and Miyake, 2005), and text

that incorporates headers and other reader-friendly attributes will be more effective than text that doesn't (Shriver,)

What we know about perception and use of value measures. As summarized elsewhere in the report, value measures are required or useful in a variety of regulatory and non-regulatory policy contexts, ranging from local government assessment and prioritization of environmental actions, to educational outreach, to federal assessment of agency programs. In some cases monetization is required, whereas in others (e.g., educational outreach by regional partnerships), narratives and visual representations of values appear to play an important role. There is little direct evidence about how such value measures are perceived, although there is considerable indirect evidence regarding their use. For example, measures that are not quantified and monetized in regulatory impact analyses appear unlikely to be fully considered or used in cost-benefit judgments. In contrast, participative decision making exercises can and do use ecological indicators as a basis for prioritizing and trading off actions to protection ecological systems and services, without monetization (ref from Denny, recent state exercises).

Text Box 8: Table of Examples of Measures from Different Examples Discussed in this Report
[To be completed with values from current draft]

Value	Measure	Characteristics	Context/use	Reference
	Willingness to pay, in $		Regulatory impact analysis	

8.2.3 Guidelines for design choices: audience assessment, user needs, and visual and interactive communication strategies

The potential interested parties for values (VPESS) include community members, policy makers, and scientists, especially environmental policy scientists. There is likely a broad public audience interested in better understanding the value of protecting ecological systems and services, but also an intermediate group of those who would use data and models, who through their analyses and activities serve as important mediators for this kind of information. (AB: refer to the assessment of audiences for values elsewhere in this VPESS report, if there is one?) They will need to access technical details and models, as well as resulting value estimates.

Effective values communication requires systematically supporting interactions with interested parties, the character of which will differ depending on the technical expertise and focus of the interested parties. In general, interactive (participative) processes are critical for improving understanding, although messages or reports (such as the ROE) are also important, especially in the context of assessment. Even for reports and indicators, having an empirical, end-user-informed basis for the design is critical. End-user engagement is itself an example of a participative process, requiring due consideration of such issues as sampling and representation. Stakeholders are likely to vary considerably in their interests, abilities, and resources such as time or access to experts who can answer technical questions. While verbal quantifiers (e.g., "many" or "very likely") are often proposed as a way of making technical information more accessible, the wide variability with which these are interpreted (Budescu and Wallsten, 1995) makes it critical to make the underlying numerical information readily available.

Appropriate use of graphical and visual approaches including GIS, can aid interpretation of quantitative information. MacEachren (1995) emphasizes the function of visualization in facilitating viewers' new and suprising insights.

Interactive communications are likely to be more effective in many circumstances than static displays. They allow users to manipulate the data or representations of the data – e.g., with sliders on interactive simulations. Interactive visualization has the potential to allow users to tailor displays to reflect their individual differences and questions. Even with exactly the same presentation, because of differences in educational or cultural background, and different intellectual abilities, people's understandings of presentation content vary. Interactive exploration tools give the audience a chance to investigate freely the part that they are either interested in or about which they still have questions.

As argued by Strecher, Greenwood, Wang, & Dumont, (1999), the advantage of interactivity lies in: a) allowance for active, instead of passive, participation of audience; b) the ability to tailor information for individual users; c) the ability to assist the assessment process; and d) the ability to visualize possible risks under different hypothesized conditions (allow users to ask "what if" questions). (Cliburn, Feddema,

Miller, & Slocum, 2002) also suggest that interactivity is a good solution if the users are overwhelmed by the complexity of the visualization. However, interactive visualization poses challenges as well. According to Encarnacao et al. (1994), interactivity is necessitated and challenged at the same time by 3-D visualization, which has become increasingly popular in visualization practice.

Finally, fundamental guidelines for risk and technical communication are generally applicable to values communication. Two examples of such guidelines are the communication principles from EPA's Risk Characterization handbook, and Guidelines for effective websites from Spyridakis. The Risk Characterization handbook principles include Transparency, Clarity, Consistency and Reasonableness. Guidelines for effective websites are available from Spyridakis (2000).

8.2.4 Communicating Uncertainty in Ecological Valuations

In order to assess how much confidence to attribute to the projections involved in the valuation, decision makers must also be informed about the analyst's own judgment of the uncertainty of the valuation and its prior steps, and the assumptions underlying the valuation analysis. Making decision makers aware of these assumptions is also important because decision makers often have to explain and justify their decisions by clarifying the assumptions driving the analysis.

In order to convey to policy makers the degree of uncertainty in an ecological valuation, the simplest expressions - whether quantitative (measures of dispersion, such as variance) or qualitative (such terms as "likely," "very likely," etc.) - are typically inadequate. Analysts can specify the central tendency of an estimate (mean or median value, as appropriate) plus a confidence interval (for example, the 95% confidence interval), but in some cases this may require possibly arbitrary judgments on the part of the analyst (Moss & Schneider 2000). Furthermore, providing policy makers with such ranges of results can be highly misleading, because those without training in probability and statistics may be likely to assume - in effect - that the probability distribution of values between the end-points is uniform, which is rarely, if ever, the case. Sensitivity analysis can help in this regard, although what is really needed is a description - verbal or pictorial - of the full probability distribution.

Institutional obstacles to conveying uncertainty may be related to the understandable reluctance of analysts to expose themselves and their work to the risk of appearing to be lacking in rigor. Analysts may thus have an unfortunate incentive to exclude or otherwise downplay components of their analyses that they fear may jeopardize the credibility of their overall effort. Suppressing less certain information runs counter to the need for transparency and the reality that all estimates have some degree of uncertainty (Arrow et al. 1996).

Historically, efforts to address uncertainty in ecological valuations - and more broadly, in Regulatory Impact Analyses (RIAs) - have been limited. But guidance set forth in the U.S. Office of Management and Budget's (OMB) Circular A-4 on Regulatory Analysis in 2003 has the potential to enhance the information provided in RIAs regarding uncertainty.

In the past, point estimates have been given far greater prominence in RIAs and other government valuations than discussions of uncertainty associated with them. Uncertainty assessments are often relegated to appendices and discussed in a manner that makes it difficult for readers to discern their significance. This is perhaps inevitable given that single point estimates can be communicated more easily than lengthy qualitative assessments of uncertainty or a series of sensitivity analyses. The ability of Monte Carlo analysis to produce quantitative probability distributions provides a means of summarizing uncertainty that can be communicated nearly as concisely as point estimates. The need for and means of communicating uncertainty in such a fashion has been addressed in the existing literature. If a summary of uncertainty in an estimate is not given prominence relative to the estimate itself, context for interpreting the estimate and opportunities to learn from uncertainty associated with it may be lost.

Some resistance to the use of formal uncertainty assessments such as through Monte Carlo analysis and prominent presentation of the results may be due to the perception that such analysis requires more expert judgment and therefore makes the results presented more speculative. Also, some might argue that, given the inevitably incomplete nature of any uncertainty analysis, prominently presenting its results would incorrectly lead readers to conclude that results of an ecological valuation are more

certain than they are. Both concerns seem to be unfounded. First, as described above, developing characterizations of uncertainty (such as for inputs in a Monte Carlo analysis) often simply involves making explicit and transparent expert judgments that necessarily already must be made to develop point estimates for those inputs. Moreover, to the extent that an uncertainty analysis is thought to be incomplete in its characterization of uncertainty, that fact can surely be communicated qualitatively. Finally, MacEachren et al. (2005) suggest animation as an effective technique for conveying uncertainties in space-time processes, which can help viewers distinguish between spatial and temporal uncertainties.

8.2.5 Evaluation

In general, it is difficult to predict the effects of communications. Good communications practice requires formative evaluation of communications as part of the design process. Summative evaluation after the fact will enable assessments of effectiveness, and continued improvement (e.g. Scriven, 1967; add Rossi and other refs)

9. RECOMMENDATIONS

Text to be added after teleconference discussions

9.1. Research

9.2. Guidance documents

9.3. Institutional Recommendations

PART 3: Methods for Implementing Approach

1. INTRODUCTION

This part of the report provides more detailed information about the methods that the committee examined for possible use in implementing the integrated and expanded valuation process proposed in Part 1. These methods differ in a number of respects, including the underlying premises and assumptions, the types of values they seek to characterize, the empirical and analytical techniques used to apply them, their data needs (inputs) and the metrics they generate (outputs), the extent to which they involve the public or stakeholders, the role that they might play in ecological valuation in different decision contexts, and the extent to which they have previously been used (by EPA or others) for this purpose. While there is no perfect way to categorize or group these methods, the Committee has organized the discussion of methods around groupings based primarily on the basic premises that underlie the different methods. In each case, the goal is not to provide an exhaustive treatise on a method; rather, it is to provide the reader with sufficient information about the methods to allow a preliminary assessment of the role that various methods could play in implementing the proposed valuation process (including strengths and possible weaknesses of different methods) and to direct the interested reader to the relevant scientific literature for further information.

2. BIO-PHYSICAL RANKING METHODS

2.1. Introduction

Text needed

2.2. Conservation Value Method

Every landscape can be characterized by a suite of ecological properties that form the basis for all environmental, social and economic values. These ecosystem and ecosystem services properties include biotic and abiotic factors that are spatially distributed across all landscapes. Biotic features include factors describing the distribution and abundance of plant and animal species, along with repeating patterns of biological associations. Abiotic factors properties include factors like soils, geology, hydrology, climate, topography, and landscape position. The combination of these bio-physical factors results in a spatially explicit map of ecological features and condition, which can include ecosystems services.

The combined values for each factor results in the continuous aggregate spatial representation of bio-physical values across the landscape. Different stakeholders may apply different values to each factor depending upon the objectives of their analysis. A conservation biologist, for example, may assign high conservation value to specific locations of viable populations of rare species and remaining stands of historic vegetation types. A wetlands biologist may assign high conservation value to functional occurrences of existing wetlands and to other landscape sites where wetlands could be restored. All bio-physical ranking exercises use one or more of these ecosystem factors to create a value surface across the landscape at appropriate scales of resolution. Each valuation exercise must be tailored to address a specific decision outcome and are documented as weighted expert judgments. All relevant ecosystem values are integrated to create a value surface across the landscape that provides a robust and transparent scientific basis for making ecosystem conservation, resource management, impact assessment, monitoring, and mitigation decisions.

These bio-physical methods produce sets of spatially explicit values that are independent from the economic value of these resources. The products from these methods represent a non-monetary value which is developed through the assessment of rarity, persistence, trends, and other indicators of each factor's contribution to sustained ecosystem diversity and integrity.

Most of these ecosystem factors also have direct resource value and provide many ecological services, the incorporation of these economic values are often integrated with the bio-physical values to inform specific decision-making processes. For example, a public agency may apply economic analyses to prioritize action between multiple options for wetland protection or impact mitigation..

Brief description of the method. The Conservation Value Method is based on a defined set of scientific criteria that supports the mapping and assessment of biological and ecological values within a regional context. This method results in a spatial representation of the stakeholder weighted value of ecological factors and ecosystem services. This method is composed of structured methodological steps which incorporate the input of stakeholder values at multiple points in the process. These analyses can be completed with current GIS-based technologies. It allows for an effective approach with multiple stakeholders to prioritize parcels in the landscape for acquisition and conservation. In some applications, this approach has been referred to as "green printing", weighting and mapping the resources that represent what people want to preserve. Groups such as Trust for Public Lands use this to focus watershed stakeholder groups on steps to implement conservation. The methodology is detailed in Grossman and Comer (2004).

This method provides a scientifically based approach to assign a value to the representative ecological factors within a specific region. These values are generated from system attributes for uniqueness, irreplaceability, level of imperilment, and ecological services.

Key Steps in the Method and Key Inputs:

a) Identify the ecological entities that will be included in the valuation process
b) Develop the "occurrence" definitions for each entity
c) Define the quality and viability values for occurrences of each entity
d) Assess the range-wide status for each entity
e) Generate an 'ecological value layer' for each entity that incorporates the values and goals of the stakeholders
f) Generate a composite spatial representation of all entities to represent stakeholder values and goals
g) Prioritize conservation and management activities through incorporation of threat and opportunity attributes

Illustration of how the method could be used as part of the CVPESS expanded and integrated framework. The method could be used in steps 1 and 3 of the C-VPESS framework.

The Conservation Value Method begins with the documentation of the intrinsic ecological values that are represented across any landscape. These values are computed through the evaluation of rarity, representation, threat, and other inherent natural factors. The value of these factors can be further modified to represent the social value of ecosystem services represented by these ecological factors. The integration of intrinsic ecosystem values with social values creates a transparent methodology for the prioritization of conservation and management actions. By definition, there is a requirement for participation of natural scientists and social scientists to complete these valuations with input from critical stakeholders at defined steps in the process.

Is value monetized or not? The primary valuation of ecological factors in the Conservation Value Method is not based on monetary value. The method can incorporate additional inputs, many of which may carry monetary attributes. As such, it has the ability to incorporate secondary monetary factors in the analyses, and can represent these results.

Does the method help to identify ecosystem service impacts of EPA action or use ecosystem service information? The Conservation Value Method can incorporate ecosystem service values and information when they are available. The method can provide a baseline of ecosystem service value for discrete landscape features and quantify the impacts of any actions (including those by EPA) from this baseline.

A methodology that is based in a GIS environment can reflect multiple decision models by adding different stakeholder values to the same landscape model. For example, one could quantify BLM land for its value as recreational use, natural resource extraction (timber, mineral, oil and gas), and water quality (denitrification, water purification) and quantity (flood control, snow pack).

Does the method identify changes in ecosystems and their services that are likely to be of greatest concern to people? The Conservation Value Method is primarily focused on providing a transparent scientific representation of ecological factors across a specific area. This information provides a critical framework to base most planning, assessment and resource management actions and decisions. In some cases, the baseline ecological integrity value itself provides the greatest utility for decision making. In other cases, this information provides the baseline map of

ecosystem factors that can project changes in ecosystems and their services that are of concern to people.

For the Conservation Value Method to be truly effective at a regional or local level it must be scientifically grounded in ecology, and be able to incorporate social and economic factors. The map of ecological value is needed to protect biodiversity while providing a knowledgeable framework for making difficult social decisions. The Agency should explore the integration of the Conservation Value Method with Social-Psychological surveys and other facilitated dialogues for experts and non-experts.

Does the method address ecological impacts that wouldn't normally be included by EPA or are not easily addressed by monetized approaches? This method provides a robust representation of biodiversity and integrity within a historic and regional context, and integrates information of specific ecosystem values to key stakeholders. This approach has broad untapped utility that will dramatically improve traditional approaches to impact analysis. It goes well beyond the ability of monetized approaches in characterizing the broader values of the ecosystem, and the impact of different change scenarios on that ecosystem.

This approach provides a spatial framework for monetized assessments as well as non-monetized endpoints. This method provides a valuable environment to address competing stakeholder values and impacts.

If this method does not present values in monetizable economic units; does it present some basis or rationale for judging or characterizing the importance of alternative expressions of value? The Conservation Value Method presents a transparent and broadly accepted methodology to spatially represent the occurrence and condition of different ecosystem factors. This spatial representation is meant to characterize alternative expressions of value across the area(s) of analysis.

Does the method involve interdisciplinary collaboration among physical/biological and social scientists; and ecologist at an early stage? The Conservation Value Method involves very detailed interactions between the physical, biological, and ecological scientists at an early stage. Shortly thereafter, it incorporates social scientists input to represent the social value attributes to these bio-eco-physical entities.

Does the method involve the public in identifying values at an early stage and have recurring involvement through the process? The Conservation Value Method incorporates

public input at multiple points in the process to ensure a high level of understanding and endorsement. This method has been developed and applied to address public planning challenges, and to be successful must incorporate public value into the analysis and public choice regarding the different action alternatives.

Are there types of data, and information that could be generated through use of the method that might be useful to have in a data/model bank so that valuation can be done more easily by EPA in the future? The Conservation Value Method provides critical information for EPA decision making and the underlying data should be based in an accessible data bank to ensure efficient access and regular use.

Status as a method. The Conservation Value Method approach represents a sequence of iterative steps that have been developed by the scientific community over the past thirty years. This methodology represents the current state of knowledge and confluence between ecosystem classification, biological conservation, and protected areas initiatives – and incorporates best practices in the social science of stakeholder surveys. These initiatives, combined with the growing body of scientific knowledge and information, have supported the creation of robust methods, datasets, tools and analyses that allow us to classify, valuate and map biological and ecological attributes across the landscape. This is helping stakeholders across multiple sectors to make better decisions regarding the conservation and management of biological and ecological resources, and the protection of ecosystem values and services.

There is widespread use of various components of these methods across US federal agencies, though the utility use of the comprehensive integrated methodology has only recently become accessible and manageable for the non-specialist. The ranking methodologies for conservation elements (plant, animals and ecosystems) has been documented in the scientific literature over many years and is in common use by numerous federal agencies (e.g., USDA FS, USFWS, NPS, and BLM). The viability and quality ranking criteria for the occurrences of conservation elements has been the topic of widespread analysis by IUCN, TNC, NatureServe and others. The conservation planning methods have emerged from Australian natural resource agencies (e.g., CSIRO) and are well published in the conservation science literature. EPA has used different components of this methodology to identify and prioritize rare and threatened species that need protection (e.g. working with the pesticide industry to protect biological diversity) and to characterize different wetland ecosystems to prioritize protection activities.

This methodology is increasingly being used by the larger planning community for different purposes at multiple scales. The examples listed below will illustrate the breadth of these applications. The Land Trust of Napa County has used the methodology to identify priority conservation acquisitions for the next ten years. The USFS is testing its use for the development and monitoring of National Forest plans. The Conservation Trust of Puerto Rico has applied these methods to clarify conservation and development priorities and options across the island. The state of Mata Grosso in Brazil is using this approach to integrate a conservation reserve program into private landholdings.

In summary, these methods are increasingly being used as the ecological framework to inform management and conservation decisions.

Table 7 provides a detailed description of standards, definitions, and use of this method outside EPA.

Table 7: Steps to Execute the Conservation Value Method and Application of the Steps Outside the EPA

STEPS TO EXECUTE THE METHOD	APPLICATION OF THE METHOD AND DIFFERENT WAYS OF USING THIS METHOD OUTSIDE THE EPA
Identify the ecological entities that will be included in the valuation process Determine the level of ecological organization on which to evaluate • For regional scale valuation, species, natural communities and ecosystems are generally used as the targets of conservation assessment and biodiversity valuation. Define and use a combination of coarse and fine conservation filter elements. Fine filter elements represent those components of biodiversity that can become extinct due to lack of knowledge or attention • Include imperiled, declining, endemic, vulnerable, "umbrella" species and subspecies, as well as Focal Communities such as unique environments, rare plant communities, rare aquatic habitats, vulnerable species aggregations, migratory stopover points, and others. Coarse filter elements are comprised of the broad vegetation types, habitats and ecological systems that represent aggregations of communities and natural landscape patterns and processes at scales useful for management and monitoring.	The valuation of fine and coarse filter conservation targets across any landscape requires a standardized approach to their definition and classification. All species and ecological classification units should apply documented classification and ranking standards. NatureServe and the network of state Heritage Program currently maintain this level of currency and standardization for over 30,000 animal taxa, 56,000 plant taxa, 7,000 vegetation types and 1,500 ecological systems. ITIS (Integrated Taxonomic Information System) similarly maintains a standardized list for species for use by scientists and federal agencies. Most fine filter targets and many coarse filter conservation elements across North America have been evaluated and given a conservation status rank.
Develop the occurrence definitions for each entity Element Occurrences document an area of land or water where the species or natural community is, or was historically present.	Biologists and ecologists have developed criteria and have been conducting inventories for many decades to document the best occurrences of these elements of conservation across the landscape. NatureServe databases alone manage and distribute information on

	the occurrences of over 500,000 imperiled species across the United States. This number grows dramatically when adding freshwater and coastal habitats, vegetation types and ecological systems.
Define the quality and viability values for occurrences of each entity Assign a relative rank (A, B, C, D) to each element occurrence according to: • Size: quantitative measure of area of occupancy, population abundance, population density, and population fluctuation. • Condition: biotic/abiotic factors, structures, processes within the occurrence as measured by population reproduction and health, development and maturity, ecological processes, species composition and biological structure, along with abiotic physical and chemical factors. • Landscape Integrity: qualitative measure of biotic factors, abiotic factors, and processes surrounding the EO. include landscape structure and extent, community development and maturity, intactness of ecological processes, species composition and biological structure, and additional abiotic physical and chemical factors. In some cases these factors can be directly measured, while in other they may be inferred/estimated indirectly	Many of coarse and fine filter occurrence quality metrics have been developed and used to provide a quality/integrity attribute to all occurrences. The quality ranks portray what experts determine to be within acceptable ranges of variation. These ranges are developed through the characterization of multiple, apparently undisturbed examples, examination of impact and response to human-induced alterations, review of literature and historical records, and the development and testing of ecological simulation models. Meaning of ranks: • "A" within the preferred ecological integrity threshold • "B" have one key factor within its acceptable range of variation • "C" do not have any key factors with their acceptable range of variation, but they are still considered to be 'restorable' • "D" no longer restorable
Assess the range-wide status for each entity This is primarily completed and most useful at global scale, but can also be applied at national, sub-national and local scales Factors for evaluation of species: total number and condition of	Global (G) Ranks have been defined as follows for : G1 – Critically imperiled: At very high risk of extinction due to extreme rarity (often 5 or fewer

occurrences (e.g., populations); population size; range extent and area of occupancy; short- and long-term trends in the above factors; scope, severity, and immediacy of threats; number of protected and managed occurrences; intrinsic vulnerability and environmental specificity. Factors for evaluation of ecological communities: • Primary factors: total number of occurrences (e.g., forest stands) and the total acreage occupied by the community. Secondary factors: geographic range over which the community occurs, long-term trends across this range, short-term trend (i.e., threats), degree of site/environmental specificity exhibited by the community, and the imperilment or rarity across the range as indicated by sub-national ranks assigned by local natural heritage programs.	populations), very steep declines, or other factors. G2 – Imperiled: At high risk of extinction due to very restricted range, very few populations (often 20 or fewer), steep declines, or other factors G3 – Vulnerable: At moderate risk of extinction due to a restricted range, relatively few populations (often 80 or fewer), recent and widespread declines, or other factors G4 – Uncommon but apparently secure: Uncommon but not rare; some cause for long-term concern due to declines or other factors G5 – Widespread, abundant and secure: Common; widespread and abundant
Generate an 'ecological value layer' for each entity that incorporates the values and goals of the stakeholders Develop conservation value surface layer for each individual filter element using the information described above for the value attributes that have been created for the global range-wide conservation status and the quality of viable occurrences.	This layer portrays each of the occurrences as well as the quality and confidence of each occurrence. The spatial portrayal of element occurrences is derived from imagery, maps and field points, along with modeled distributions of specific elements.
Generate a composite spatial representation of all entities that represents stakeholder values and goals Combine 'conservation value layers' for selected elements across a planning or assessment jurisdiction to create an aggregated 'conservation value summary' that provides a spatially explicit representation of the biodiversity and conservation values that are important to the conservation and resource management community. Different user groups can select the types of elements that they need to assess across the jurisdiction, and they can also modify the relative	

conservation weight of each fine and coarse filter conservation element. This will provide a customized conservation surface that portrays the values that they will need to incorporate into their planning and assessment work. This also becomes a baseline for monitoring the effects of their programs to manage for biodiversity value over time.	
Prioritize conservation and management actions through incorporation of threat and opportunity attributes. The conservation values outlined above can be modified to reflect values that are relevant to a specific assessment. Zoning policies, growth models, economic values, ecological services and other values help to identify the effect of different or future scenarios relative to the current or desired future condition of the landscape.	

Strengths/Limitations

Conceptual Strengths/Limitations: The Conservation Value Method will create a quantitative spatial representation of ecological and biological values within a regional context. The spatial range of these analyses can vary from local to regional scales. This data provides a baseline for a broad range of natural resource assessment and management decisions, and can be integrated with spatial monetary valuations to inform cost-effective land management and regulatory decisions. The specific decisions will determine that types of data and analyses that are required to address the question.

Decision contexts where this method could be used by EPA include:

- Enumeration of biodiversity protection implications that result from policy changes (i.e., change of protection status for isolated wetlands).
- Identification of critical riparian habitat
- Prioritization of remediation action on superfund sites
- Due diligence reviews and Environmental Impact Statements as a prerequisite for permitting.
- Identification of reference conditions for establishment of baseline quality metrics for wetland and aquatic habitats.
- Assessment of the status of target species and ecosystems.
- Analysis of mitigation equivalencies and priorities
- Baseline information for ecosystem integrity and environmental impact monitoring

The method's strengths:

- The method is adaptable to address different questions
- The method can be run repeatedly to represent temporal change or different landscape scenarios.
- Results are commonly aggregated to derive a single benefits number, but all of the native data is constantly maintained in the system and can be presented separately.

- The output is both understandable and communicable to the interested audience and other stakeholders. Provides the opportunity for visualization of outcomes that many other methods lack.
- The results are repeatable, and the process and algorithms are very transparent.

The method's weaknesses: Issues with the lack of data, the currency and confidence in available data, along with access to 'sensitive' data represent potential obstacles for the application of this method. There are many ways to create surrogate datasets that will allow users to adapt to different types of 'barriers'. Some training and tools are also required to use this method.

Practical Strengths/Limitations:

- The assumption is that there is sufficient coverage of standardized biodiversity data required to implement these methods. The standards for each step of the method have been developed, and the data that is required will be dependent upon the specific application questions. Where sufficient data does not yet exist, additional resources will need to develop this information in order to complete the methodology. In some cases, surrogate information and models are required to incorporate the spatial representation of poorly inventoried conservation targets across the landscape..
- This method requires local scientific data, knowledgeable scientific interpretation and conservation planning expertise. The magnitude of the need is contingent upon the application and the current state of data and knowledge. There are many sources available from which to obtain this knowledge.

Treatment of Uncertainty: There are confidence measures built into the methodology that can be integrated into the decision making analysis or displayed independently for consideration. The most significant sources of uncertainty in the use of this method include:

- The variability in the quantity and quality of the data.

- The limitations of scientific understanding of distribution and quality criteria for some ecological factors.
- The level of stakeholder understanding of the linkages between ecological components and the services they value.

Research needs: There is both a need and an opportunity to actively explore integration of stakeholder elicitation approaches (e.g. Social Scientific surveys) with ecological condition mapping. Additional R&D to show how GIS based systems could be designed to integrate monetized and other quantitative valuation approaches on a common spatial and temporal GIS background could yield significant benefits.

Key References

Brown, N., L. Master, D. Faber-Langendoen, P. Comer, K. Maybury, M. Robles, J. Nichols, and T. B. Wigley. 2004. Managing Elements of Biodiversity in Sustainable Forestry Programs: Status and Utility of NatureServe's Information Resources to Forest Managers. National Council for Air and Stream Improvement Technical Bulletin Number 0885.

Grossman, D.H. and P.J. Comer. 2004. Setting Priorities for Biodiversity Conservation in Puerto Rico. NatureServe Technical Report.

Riordan, R. and K. Barker. 2003. Cultivating biodiversity in Napa. Geospatial Solutions.

Stoms, D. M., P. J. Comer, P. J. Crist and D. H. Grossman. 2005. Choosing surrogates for biodiversity conservation in complex planning environments. Journal of Conservation Planning 1.

2.3. Rankings Based on Energy and Material Flows

Energy and material flow analysis is the quantification of the flows of energy and materials through complex ecological and/or economic systems. A recent National Academy report covers the basic elements and need for such analyses (Committee on Materials Flows Accounting of Natural Resources, Products and Residuals, 2004). These analyses are based on an application of the first (conservation of mass and energy) and second (entropy) laws of thermodynamics to ecological-economic systems. These methods are covered here because they typically reveal insights not available through other means. In addition, they hold the possibility of treating ecological and economic systems in the same conceptual framework.

This section provides general background on energy and material flow accounting followed by three specific applications which attempt to use this method for ranking alternatives: embodied energy, emergy, and ecological footprint.

Both the embodied energy and ecological footprint methods use input-output analysis or flow accounting methods. Embodied energy analyses estimates the direct and indirect energy (or more correctly available energy or "exergy") cost of goods and services. Cleveland (1987) traced the early roots of this work dating back to the Physiocrats. Ecological footprint analysis estimates the biologically productive land area required (directly and indirectly) to meet various consumption patterns. Emergy analysis shares many of the goals of embodied energy analysis, but does not use input-output methods. Instead, it uses its own idiosyncratic methods to derive estimates of the energy cost of goods and services.

Although there is no stated Agency policy to use or develop supplemental valuation methodologies in this area, there is substantial Agency interest in how Energy and Material Flow methods might aid decision-making. Recent efforts to explore the utility of such methods, mostly at the regional or local level, have been published (Bastianoni et. al. 2005, Campbell, 2001, 2004; Lu et. al, 2006) and others are underway.

<u>Background on the approach:</u> The energy and environmental events of the 1960's and 1970's prompted a number of economists, ecologists and physicists to examine the energy and material flows underlying the economic process. Economists such as Boulding (1966) and Georgescu-Roegen (1971, 1973) demonstrated the environmental and economic

implications of the mass and energy balance principle. Ecologists such as Lotka (1922) and Odum (Odum and Pinkerton 1955, Odum 1971) pointed out the importance of energy in the structure and evolutionary dynamics of ecological and economic systems. And physicists such as the Nobel Laureate Prigogine contributed substantially to our understanding of the irreversible nature of thermodynamic processes.

In thermodynamics terminology there are three categories of systems: open, closed, and isolated. Open systems allow matter and energy to cross the boundaries. Closed systems allow only energy to cross the boundaries (i.e. closed to matter, but not energy). Isolated systems allow nothing to cross the boundaries. The second law (that entropy always increases) applies only to isolated systems. The earth is (for the most part) a closed system (not an isolated system) with lots of energy crossing the boundaries, which is why it will not run down (or at least it can compensate for the running down with new energy inputs).

The first law of thermodynamics (conservation of mass and energy) has formed the basis for a number of important contributions. The integration of this law with the economic system was first made explicit in the context of an economic general equilibrium model by Ayres and Kneese (1969) and subsequently by Mäler (1974), but it is also a feature of a series of linear models developed after 1966 (Cumberland 1966, Victor 1972, Lipnowski 1976). All reflect the recognition that a closed physical system must satisfy the conservation of mass condition, and hence that economic growth necessarily increases both the extraction of environmental resources and the volume of waste deposited in the environment.

Ayres (1978) described some of the important implications of the laws of thermodynamics for the production process, including the limits they place on the substitution of human-made capital for natural capital. Indeed a primary insight of this work is that natural capital and human-made capital ultimately are *complements* because both manufactured and human capital require materials and energy for their own production and maintenance (Costanza 1980). This insight has implications for the interpretation of traditional economic production functions.

Furthermore, work based on the laws of thermodynamics shows the limits of technological progress. In sectors that are largely concerned with processing and/or fabricating materials, technical change is subject to diminishing returns as it approaches a thermodynamic minimum (Ayres 1978), i.e. the minimum amount of energy and materials

required to produce a unit of output. Ruth (1995) used equilibrium and non-equilibrium thermodynamics to describe the materials-energy-information relationship in the biosphere and in economic systems. In addition to illuminating the boundaries for material and energy conversions in economic systems, thermodynamic assessments can provide insights about depletion and degradation of the natural world, phenomena that often escape traditional economic analyses based on supply and demand curves.

Biophysical foundations have been incorporated into models of natural resource supply and of the relationship between energy use and economic performance. Cleveland and Kaufmann (1991) developed econometric models that explicitly represent and integrate the geologic, economic, and political forces that determine the supply of oil in the United States. For explaining the historical record, these models are superior to any from a single discipline. Larsson et al. (1994) also used energy and material flows to demonstrate the dependence of a renewable resource such as commercial shrimp farming on the services generated by marine and agricultural ecosystems.

One important advance generated by this work is recognition of the economic importance of energy quality, namely, that a kcal of primary electricity can produce more output than a kcal of oil, a kcal of oil can produce more output than a kcal of coal, and so on. Solar energy, although ubiquitous, requires man-made energy and materials to capture and upgrade it to more usable forms. Hence, the relative abundance and universality of solar energy does not translate into usable quality and abundant availability in the right forms at the right locations. An analogy can be made with water and water quality: "Water, water, everywhere, but not a drop to drink." Odum (1971) describes how energy use in ecological and economic hierarchies tends to increase the quality of energy, and that significant amounts of energy are dissipated to produce higher quality forms that perform critical control and feedback functions which enhance the survival of the system. Cleveland et al. (1984) and Kaufmann (1992) show that much of the decline in the energy/real GDP ratio in industrial nations is due to the shift from coal to petroleum and primary electricity. Their results show that autonomous energy-saving technical change has had little, if any, effect on the energy/real GDP ratio. Stern (1993) finds that accounting for fuel quality produces an unambiguous causal connection between energy use and economic growth in the United States, confirming the unique, critical role that energy plays in the production of wealth.

The analysis of energy flows has also been used to illuminate the structure of ecosystems (e.g. Odum, 1957). Hannon (1973) applied input-output analysis (originally developed to study interdependence in economies) to the analysis of energy flow in ecosystems. This approach quantifies the direct plus indirect energy that connects an ecosystem component to the remainder of the ecosystem. Hannon demonstrates this methodology using energy flow data from the classic study of the Silver Springs, Florida food web (Odum, 1957) and using this framework to estimate "shadow prices" of ecosystem goods and services (Hannon et al. 1986, 1991, Costanza and Hannon 1989).

Background on the approach. The energy and environmental events of the 1960's and 1970's (including the Arab oil embargos of 1973 and 1979) prompted a number of economists, ecologists and physicists to examine the energy and material flows underlying the economic process. Economists such as Boulding (1966) and Georgescu-Roegen (1971, 1973) demonstrated the environmental and economic implications of the mass and energy balance principle. Ecologists such as Lotka (1922) and Odum (Odum and Pinkerton 1955, Odum 1971) pointed out the importance of energy in the structure and evolutionary dynamics of ecological and economic systems. And physicists such as the Nobel Laureate Prigogine contributed substantially to our understanding of the irreversible nature of thermodynamic processes.

In thermodynamics terminology there are three categories of systems: open, closed, and isolated. Open systems allow matter and energy to cross the boundaries. Closed systems allow only energy to cross the boundaries (i.e. closed to matter, but not energy). Isolated systems allow nothing to cross the boundaries. The second law (that entropy always increases) applies only to isolated systems. The earth is (for the most part) a closed system (not an isolated system) with lots of energy crossing the boundaries, which is why it will not run down (or at least it can compensate for the running down with new energy inputs).

The first law of thermodynamics (conservation of mass and energy) has formed the basis for a number of important contributions. The integration of this law with the economic system was first made explicit in the context of an economic general equilibrium model by Ayres and Kneese (1969) and subsequently by Mäler (1974), but it is also a feature of a series of linear models developed after 1966 (Cumberland 1966, Victor 1972, Lipnowski 1976). All reflect the recognition that a closed physical system must satisfy the conservation

of mass condition, and hence that economic growth necessarily increases both the extraction of environmental resources and the mass of waste deposited in the environment.

Ayres (1978) described some of the important implications of the laws of thermodynamics for the production process, including the limits they place on the substitution of human-made capital for natural capital. Indeed a primary insight of this work is that natural capital and human-made capital ultimately are complements because both manufactured and human capital require materials and energy for their own production and maintenance (Costanza 1980). This insight has implications for the interpretation of traditional economic production functions.

Furthermore, work based on the laws of thermodynamics shows the limits of technological progress. In sectors that are largely concerned with processing and/or fabricating materials, technical change is subject to diminishing returns as it approaches a thermodynamic minimum (Ayres 1978), i.e. the minimum amount of energy and materials required to produce a unit of output. Ruth (1995) used equilibrium and non-equilibrium thermodynamics to describe the materials-energy-information relationship in the biosphere and in economic systems. In addition to illuminating the boundaries for material and energy conversions in economic systems, thermodynamic assessments can provide insights about depletion and degradation of the natural world, phenomena that often escape traditional economic analyses based on supply and demand curves.

Biophysical foundations have been incorporated into models of natural resource supply and of the relationship between energy use and economic performance. Cleveland and Kaufmann (1991) developed econometric models that explicitly represent and integrate the geologic, economic, and political forces that determine the supply of oil in the United States. For explaining the historical record, these models are superior to any from a single discipline. Larsson et al. (1994) also used energy and material flows to demonstrate the dependence of a renewable resource such as commercial shrimp farming on the services generated by marine and agricultural ecosystems.

One important advance generated by this work is recognition of the economic importance of energy quality, namely, that a kcal of primary electricity can produce more output than a kcal of oil, a kcal of oil can produce more output than a kcal of coal, and so on. Solar energy, although ubiquitous, requires man-made energy and materials to capture and

upgrade it to more usable forms. Hence, the relative abundance and universality of solar energy does not translate into usable quality and abundant availability in the right forms at the right locations. An analogy can be made with water and water quality: "Water, water, everywhere, but not a drop to drink." Odum (1971) describes how energy use in ecological and economic hierarchies tends to increase the quality of energy, and that significant amounts of energy are dissipated to produce higher quality forms that perform critical control and feedback functions which enhance the survival of the system. Cleveland et al. (1984) and Kaufmann (1992) show that much of the decline in the energy/real GDP ratio in industrial nations is due to the shift from coal to petroleum and primary electricity. Their results show that autonomous energy-saving technical change has had little, if any, effect on the energy/real GDP ratio. Stern (1993) finds that accounting for fuel quality produces an unambiguous causal connection between energy use and economic growth in the United States, confirming the unique, critical role that energy plays in the production of wealth.

The analysis of energy flows has also been used to illuminate the structure of ecosystems (e.g. Odum, 1957). Hannon (1973) applied input-output analysis (originally developed to study interdependence in economies) to the analysis of energy flow in ecosystems. This approach quantifies the direct plus indirect energy that connects an ecosystem component to the remainder of the ecosystem. Hannon demonstrates this methodology using energy flow data from the classic study of the Silver Springs, Florida food web (Odum, 1957) and using this framework to estimate "shadow prices" of ecosystem goods and services (Hannon et al. 1986, 1991, Costanza and Hannon 1989)

2.3.1 Embodied Energy Analysis

Brief description of the method: Ecologists and physical scientists have also proposed an "energy theory of value", either to complement or replace the standard neoclassical theory of subjective utility based value (Soddy 1922, Odum 1971, 1983, Slesser 1973, Gilliland 1975, Costanza 1980, Cleveland et al. 1984, Hall et al., 1992). It is based on thermodynamic principles where solar energy is recognized to be the only "primary" input to the global ecosystem.

This theory of value represents a return to the classical ideas of Ricardo and more recently Sraffa (1960), but with some important distinctions. The classical economists recognized that if they could identify a "primary" input to the production process then they

could explain exchange values based on production relationships. The problem was that neither labor nor any other single commodity was really "primary," since they all require each other for their production.

Costanza (1980) has pointed out that, at the global scale, the traditional "primary" factors are really "intermediate" factors of production. Available energy or exergy is the only "basic" commodity and is ultimately the only "scarce" factor of production, thereby satisfying the criteria for a production-based theory that can explain exchange values. Costanza noted that available energy has those characteristics which satisfy the criteria for a "primary" input: it is ubiquitous and essential to all production processes and cannot be replaced by something else (although substitution may occur between types of energy). An energy theory of value thus posits that, at least at the global scale, free or available energy from the sun (plus past solar energy stored as fossil fuels and residual heat from the earth's core) is the only "primary" input to economic production. Labor, manufactured capital, and natural capital are "intermediate inputs". Thus, one could base a theory of value on the use in production of available energy that avoids the problems the classical economists encountered when trying to explain exchange values in economic systems using only labor inputs.

<u>Status as a method:</u> Estimating total "energy" consumption for an economy is not a straightforward matter because (as noted above) not all fuels are of the same quality--i.e., they vary in their available energy, degree of organization, or ability to do work. Electricity, for example, is more versatile and cleaner in end use than petroleum, and it also costs more energy to produce. In an oil-fired power plant it takes from 3-5 kcal of oil to produce each kcal of electricity. Thus, adding up the raw heat equivalents of the various forms of fuel consumed by an economy without accounting for fuel quality can distort the picture, especially if the mix of fuel types is changing over time.

There have been a few attempts to empirically test this theory using both time series data and cross sectional data. Studies that have tried to adjust for fuel quality have shown a very close relationship between "available energy" consumption and economic output. Cleveland et al (1984) and Kaufmann (1992) have shown that almost all of the changes in E/GNP (or E/GDP) ratios in the US and OECD countries can be explained by changes in fuel quality and the percent of personal consumption expenditures (PCE) spent directly on fuel.

The latter effect is due to the fact that PCE is a component of GNP and spending more on fuel directly will raise GNP without changing real economic output. Much of the apparent gain in energy efficiency (decreasing E/GNP ratio) is due to shifts to higher quality fuels (like natural gas and primary electricity) from lower quality ones (like coal). Renewable energy sources are generally lower quality and shifts to them may cause significant increases in the E/GNP ratio.

Another way of looking at the relationship between available energy and economic output uses cross-sectional rather than time-series data. This avoids some of the problems associated with changes in fuel mix and distortions in GNP. For example, Costanza (1980) and Costanza and Herendeen (1984) used an 87-sector input-output model of the US economy for 1963, 1967, and 1973, modified to include households and government as endogenous sectors (to include labor and government energy costs) to investigate the relationship between direct and indirect energy consumption (embodied energy) and dollar value of output. They found that dollar value of sector output was highly correlated (R2 = .85 to .98) with embodied energy when this was calculated including the energy costs of labor, government, and environmental inputs (though not with direct energy consumption or with embodied energy calculated excluding labor and government energy costs). Thus, if one makes some necessary adjustments to estimates of energy consumption in order to better assess "available energy", it appears that the empirical link between available energy cost and economic value is rather strong.

As described above, the embodied energy method assesses the direct and indirect energy costs of economic and ecological goods and services. These methods are most applicable on a broad policy scale such as energy policy and questions of sustainability. In fact, if the United States were ever to undertake serious moves toward ecological tax reform (the replacement of taxes on labor and income with taxes on energy and environmental pollution), embodied energy calculations would be extremely useful for estimating tax rates based on energy content. For regulatory decisions that fall within EPA's discretion, energy and material flow analyses may be useful in helping to value ecosystem services, especially those services that are far removed from consumer preferences. This includes services like nutrient cycling, waste treatment, and erosion control, that revealed or stated preference methods may not be able to adequately address, because consumers are not informed about

the contribution of these services to their welfare or they have simply never "constructed" preferences for these services. In general, a "pluralistic" approach is needed to valuing ecosystem services. Comparing and contrasting the results from conventional approaches and energy-based approaches may prove quite useful, given the large uncertainties all around (see Costanza et al. 1989 for an example of wetlands valuation using this pluralistic approach).

Strengths/Limitations: Some neoclassical economists have criticized the energy theory of value as an attempt to define value independent of consumer preferences (see Heuttner, 1976). This criticism is, on the one hand, axiomatic, since a major purpose of an energy theory of value was to establish a theory of value not completely determined by individual preferences. On the other hand, techniques for calculating embodied energy utilize economic input-output tables. These tables summarize production interdependencies but they are not completely independent of consumer preferences, which helped to structure the production interdependencies over time.

In summary, the energy theory of value overcomes some of the problems with earlier production-based theories of value encountered by the classical economists and does a fairly good job of explaining exchange values empirically in the few cases where it has been tested. Despite the controversy and ongoing debate about the validity of an energy theory of value (Brown and Herendeen, 1996), it seems to be the only reasonably successful attempt to operationalize a general biophysical theory of value (see also Patterson 2002). Energy (and earlier labor) theories of value are inherently based on relative production costs. Thus it is more accurate to speak of energy cost or labor cost and not energy value or labor value. However, in economic systems it is well known that cost and price will, in general, come to equilibrium. This is the essence of the basic ideas of supply and demand. If a commodity has a much higher value than its cost of production, profits will be high and more of the commodity will be produced (with increasing marginal cost) until cost just equals price and profits are 0 for the last unit of production. Likewise, if cost is much higher than price, less will be produced until they are again equal. Therefore, a method which estimates costs, while not technically estimating price or value, should, in fact, be a fairly good approximation to price and value in those cases where markets have reached an approximate equilibrium.

Since markets are in constant flux and are far from perfect, we would expect there to be some divergence between cost and price in real systems. The question then becomes: "are the commodities overpriced or underpriced?" Should the energy costs be taken as the standard and market exchange values be adjusted or should the energy costs be ignored and the market prices taken as the standard? Given, on the one hand the enormous data requirements to calculate energy costs accurately, and on the other hand the pervasive market imperfections complicating market prices, there is no unambiguous right answer to this question. But one can learn much by looking at energy costs and market prices and comparing their degree of correspondence. Costanza (1980) did just that at the aggregate level of an 87 sector Input-Output model and found a fairly high degree of correspondence for those sectors where markets were fairly functional and lacking severe externalities, and low correspondence for those sectors where they were not (basically the extractive sectors at this level of aggregation). Since the method is based on the measurement of biophysical variables, it is less prone to bias caused by subjective or "framing" issues. However, given the large data requirements and the need to aggregate sectors in some inherently subjective way, caveats and caution must accompany any conclusions.

In terms of valuing ecosystem services, at least one study has compared the embodied energy analysis approach with conventional preference-based approaches to valuation for the case of coastal wetlands in Louisiana (Costanza et al. 1989). This study found reasonable agreement between the two approaches, and concluded that there was significant benefit in performing both kinds of analyses simultaneously. Costanza et al. (1997) also included (in the supplementary information) several embodied energy analyses based methods in their synthesis of the value of global ecosystem services. While these estimates were not included in the numerical totals or averages, it was concluded that the energy analysis-based estimates "showed fairly close agreement" with the preference-based methods.

2.3.2 Emergy Analysis

Brief description of the method: Emergy analysis shares many of the same goals and assumptions as embodied energy analysis. For example, solar Emergy, is defined as "the available solar energy used up directly and indirectly to make a service or product" (Odum, 1996). Emergy analysis differs from embodied energy analysis and ecological footprint analysis in terms of the method used to estimate the energy required (directly and indirectly)

to produce a service or product. While embodied energy and footprint analysis use input-output based methods (a well-developed set of methods for this type of accounting), emergy analysis uses an idiosyncratic method.

Emergy analysis starts with the creation of an energy flow diagram. The "Solar Transformity" is then defined as " the solar emergy required to make one Joule of a service or product" (Odum, 1996). This is calculated by dividing any flow in the diagram by the total solar energy input.

Odum and coworkers have thus calculated the emergy of the earth' s main processes, such as, the total surface wind, rain water in streams, the sedimentary cycle, and waves absorbed on shore, to be that of the total emergy input to the Earth (Odum, 1996). Each of these processes is assigned the total value of incoming sunlight because they are considered co-products of the global geological cycle and cannot be produced independently with less amount of the total emergy.

One can easily see that, while this method yields numbers in profusion, the method does not obey most basic accounting principles concerning additivity

Status as a method and Strengths/Limitations: Hau and Baksi (2004) note that emergy has encountered a lot of resistance and criticism, particularly from economists, physicists, and engineers. Consequently, it has not been used much outside a small circle of researchers. They state that limited use of emergy analysis may be due to inadequate attention to details, poor communication of its potential importance, and lack of clear links with related concepts in other disciplines.

But the major reason is that emergy's idiosyncratic accounting method does not produce what it claims – an estimate of the energy cost of goods and services – in any way that is credible or useful. Because of this, emergy has been characterized as simplistic, contradictory, misleading and inaccurate (Ayres, 1998; Cleveland et al., 2000; Mansson and McGlade, 1993; Spreng, 1988).

2.3.3 Ecological Footprint Analysis

Brief description of the method: The ecological footprint (EF) method is a variation of energy and material flow analysis that converts the impacts to units of land rather than energy or dollars. The EF for a particular population is defined as the total "area of productive land and water ecosystems required to produce the resources that the population

consumes and assimilate the wastes that the population produces, wherever on Earth that land and water may be located" (Rees 2000). Input-output analysis methods (see above) are used to estimate direct and indirect land requirements.

Status as a method and Strengths/Limitations: The EF is an effective heuristic and pedagogic device for presenting current total human resource use in a way that communicates easily to a broad range of people. Although there are ongoing debates about specific methods for calculating the EF (cf. Costanza 2000, Herendeen 2000; Simmons et al. 2000) using land area as a numeraire seems to communicate well to those who have trouble with money or energy as a numeraire. There are a large number of groups utilizing the method to help communicate this information (c.f. http://www.footprintnetwork.org/). A Google search on "ecological footprint" turned up over 1 million hits, a testament to the concept's current popularity.

In terms of valuing ecosystem services, the EF concept is most useful as an index of the quantity of ecosystem services consumed (expressed in units of a standardized land area) for various consumption patterns.

Key References

Ayres, R. U., 1978. Application of physical principles to economics. pp. 37-71 in: R. U. Ayres. Resources, Environment, and Economics: Applications of the Materials/Energy Balance Principle. John Wiley and Sons, New York

Ayres, R.U., 1988. Ecology vs. Economics: Confusing Production and Consumption. Center of the Management of Environmental Resources, INSEAD, Fontainebleau, France.

Ayres, R. U. and A. V. Kneese. 1969. Production, consumption, and externalities. The American Economic Review. 59:282-297Brown, M.T. and Herendeen, R.A. 1996. "Embodied Energy Analysis and Energy Analysis: A Comparative View," Ecological Economics 19, 219-235.

Bastianoni, S., D.E. Campbell, L.Susani and E. Tiezzi. 2005, The Solar Transformity of Oil and Petroleum Natural Gas. Ecological Modeling. 186: 212 – 220.

Brown, M.T. and Ulgiati S., 1997. Emergy-based indices and ratios to evaluate sustainability: monitoring economies and technology toward environmentally sound innovation. Ecol. Eng., 9: 51-69.

Brown, M.T. and Ulgiati S., 2002. Emergy evaluations and environmental loading of electricity production systems.

Boulding, K. E. 1966. The economics of the coming spaceship Earth. pp 3-14 In: H. Jarrett (Ed.) Environmental quality in a growing economy. Resources for the Future/Johns Hopkins University Press, Baltimore, MD

Campbell, D. E. 2001. Proposal for Including What is Valuable to Ecosystems in Environmental Assessments. Environmental Science and Technology, 35: 2867-2873

Campbell, D. E. 2004. Global Transition to Sustainable Development. Proceedings of IV Biennial International Workshop" Advances in Energy Studies". UNICAMP, Campinas, SP, Brazil, June 16-19, 2004, Pages 11 - 28

Cleveland, C. J. 1987. Biophysical economics: historical perspective and current research trends. Ecological Modeling. 38:47-74.

Cleveland, C.J., Costanza, R., Hall, C.A.S and Kaufmann, R. 1984. "Energy and the U.S. Economy: a Biophysical Perspective." Science 225, 890-897.

Cleveland, C.J., Kaufmann, R.K. and Stern, D.I., 2000. Aggregation and the role of energy in the economy. Ecol. Econ., 32: 301-317.

Cleveland, C. J. and Kaufmann, R. K. 1991. Forecasting ultimate oil recovery and its rate of production: incorporating economic forces into the models of M. King Hubbert. The Energy Journal, 12:17-46.

Committee on Materials Flows Accounting of Natural Resources, Products and Residuals. 2004. Materials count: the case for material flows analysis. The National Academy Press, Washington DC. 124 pp.

Costanza, R. 1980. "Embodied energy and economic valuation," Science 210:1219-1224.

Costanza, R. (ed.) 2000. Forum: The Ecological Footprint. Special Section of Ecological Economics. Vol. 32, no. 3. Pp. 341-394

Costanza, R. 2004. Value theory and energy. Pp. 337-346 In: C. Cleveland (Ed.) Encyclopedia of Energy. vol 6 Elsevier, Amsterdam

Costanza, R. and R. A. Herendeen. 1984. Embodied Energy and Economic Value in the United States Economy: 1963, 1967, and 1972. Resources and Energy. 6;129-163.

Costanza, R., S. C. Farber, and J. Maxwell. 1989. The valuation and management of wetland ecosystems. Ecological Economics 1:335-361.

Costanza, R. and B. M. Hannon. 1989 Dealing with the "mixed units" problem in ecosystem network analysis. pp. 90-115 in: F. Wulff, J. G. Field, and K. H. Mann (eds). Network analysis of marine ecosystems: methods and applications. Coastal and Estuarine Studies Series, Springer-Verlag, Heidleberg. 284 pp.

Costanza, R., R. d'Arge, R. de Groot, S. Farber, M. Grasso, B. Hannon, S. Naeem, K. Limburg, J. Paruelo, R.V. O'Neill, R. Raskin, P. Sutton, and M. van den Belt. 1997. The value of the world's ecosystem services and natural capital. Nature 387:253-260.

Cottrell, W. F., 1955. Energy and society. McGraw-Hill, New York.

Cumberland, J.H. 1966. A regional inter-industry model for analysis of development objectives, Regional Science Association Papers 17: 65-94.

Georgescu-Roegen, N. 1971. The entropy law and the economic process. Harvard University Press, Cambridge, MA.

Georgescu-Roegen, N. 1973. The entropy law and the economic problem. pp. 49-60 in: H. E. Daly. Economics, ecology, ethics: essays toward a steady-state economy. W. H. Freeman. San Francisco. 372 pp.

Gilliland, M.W. 1975. "Energy analysis and public policy," Science 189, 1051-1056.

Hall, C.A.S., Cleveland, C.J. and Kaufmann, K. 1992. Energy and Resource Quality: The Ecology of the Economic Process. (University of Colorado Press, Colorado).

Hannon, B. 1973, The structure of ecosystems, Journal of Theoretical Biology, 41:535-46.

Hannon, B., R. Costanza, and R. A. Herendeen. 1986. Measures of energy cost and value in ecosystems. The Journal of Environmental Economics and Management. 13:391-401.

Hannon, B., R. Costanza, and R. Ulanowicz. 1991. A general accounting framework for ecological systems: a functional taxonomy for connectivist ecology. Theoretical Population Biology 40:78-104.

Hau, J.L. and B.R. Basshi, 2004 The promise and problems of emergy analysis. Ecological Modelling 2004:178(1-2) 215

Herendeen, R. 2000. Ecological footprint is a vivid indicator of indirect effects. Ecological Economics. 32:357-358

Kaufmann, R. K. 1992. A biophysical analysis of the energy/real GDP ratio: implications for substitution and technical change. Ecological Economics 6:35-56.

Lagerberg, C. and Brown, M.T., 1999. Improving agricultural sustainability: The case of Swedish greenhouse tomatoes. J. Clean. Prod., 7(6): 421-434.

Larsson, J., C. Folke, and N. Kautsky. 1994. Ecological limitations and appropriation of ecosystem support by shrimp farming in Colombia. Environmental Management. 18:663-676

Lipnowski, I.F. 1976. An Input-Output Analysis of Environmental Preservation, Journal of Environmental Economics and Management 3: 205-214.

Lotka, A. J. 1922. Contribution to the energetics of evolution. Proceedings of the National Academy of Science, 8:147-155

Lu H-F, D.E. Campbell, Z-A Li and H. Ren. 2006. Energy Synthesis of an Agro-forest Restoration System in Lower Subtropical China. Ecological Engineering 27: 175-192

Mäler, K.-G. 1974 Environmental economics - a theoretical inquiry. Johns Hopkins Press, Baltimore.

Mansson, B.A. and McGlade, J.M., 1993. Ecology, thermodynamics and H.T. Odum' s conjectures. Oecologia, 93: 582-596.

Odum, H. T. and R. C. Pinkerton. 1955. Time's speed regulator: the optimum efficiency for maximum power output in physical and biological systems. American Scientist, 43:331-343.

Odum, H. T. 1957. Trophic structure and productivity of Silver Springs, Florida. Ecological Monographs, 27, 55-112.

Odum, H.T. 1971. Environment, Power and Society. (John Wiley and Sons, New York).

Odum, H.T. 1983. Systems Ecology: An Introduction. (John Wiley and Sons, New York).

Odum, H.T., 1996. Environmental Accounting: EMERGY and environmental decision making. John Wiley, New York, 370 pp.

Odum, H.T., Brown, M.T. and Brandt-Williams, S.L., 2000. Folio #1, Introduction and Global Budget. Handbook of Emergy Evaluation, Center for Environmental Policy, Environmental Engineering Sciences, University of Florida, Gainesville.

Nicolis G. and Prigogine I. 1977. Self-organisation in non-equilibrium systems. John Wiley, New York,

Perrings, C. 1986. Conservation of mass and instability in a dynamic economy-environment system. Journal of Environmental Economics and Management. 13:199-211

Podilinsky, S. 1883. Menschliche Arbeit und Einheit der Kraft. Die Neue Zeit, Marc-April.

Prigogine I. and Stengers I. 1984. Order Out of Chaos: Man's New Dialogue with Nature, New York, Bantam Books.

Quesnay, F., 1758. Tableau Economique. In: Kuczynski, M. and Meek, R. L. (Editor), Quesnay's Tableau Economique. Macmillian, London.

Rees, W. E. Eco-footprint analysis: merits and brickbats. Ecological Economics. 32:371-374

Ruth, M. 1995. Information, order and knowledge in economic and ecological systems: implications from material and energy use. Ecological Economics 13:99-114.

Simmons, C., K. Lewis, and J. Moore. 2000. Two feet - two approaches: a component-based model of ecological footprinting. Ecological Economics. 32:375-380

Slesser, M. 1973. "Energy analysis in policy making," New Scientist. 58: 328-330.

Soddy, F., 1922. Cartesian economics. Hendersons, London.

Sraffa, Piero, 1960. Production of Commodities by Means of Commodities: Prelude to a Critique of Economic Theory. (Cambridge University Press, Cambridge).

Stern, D. I. 1993. Energy use and economic growth: a multivariate approach. Energy Economics, 15:137-150.

Victor, P. 1972. Pollution, economy and environment. George Allen and Unwin, .London.

3. SOCIO-PSYCHOLOGICAL APPROACHES

3.1. Introduction

That "we the people" should have a say in the policies and actions of our public agencies is inherent to a democratic society. Along with other public agencies EPA has a number of laws, regulations and guides that acknowledge and implement public input into agency policies and decision making. For EPA, public involvement "… means that the Agency considers public concerns, values, and preferences when making decisions" (EPA 2003, p 1). The social-psychological methods for assessing the value of protecting ecosystems and ecosystem services described in this section can be viewed as an important component of the broader public involvement program for EPA. Surveys are proven, effective methods for identifying public values and concerns ("what people care about"), for determining what environmental outcomes and associated social consequences the public prefers, accepts and supports, and for predicting how various publics are likely to respond to particular changes in environmental conditions and to alternative management means for achieving those ends. Along with focus groups and individual narrative interviews surveys can provide important inputs to and supports for agency decision making, including providing reliable and valid measures of the relative value to society of alternative ecosystem protection policies and outcomes.

While the need to consider public concerns, values and intentions is uncontroversial, accomplishing this goal can be complicated. Different segments of society may disagree on what matters most, and there can be substantial differences in people's opportunities and abilities to make their concerns and desires known (especially those yet unborn). EPA's charge to protect ecosystems and ecosystems services reflects widely shared public concerns and values (e.g., Dunlap et al. 2000), but the formulation and implementation of specific ecological protection policies involves substantial scientific and technical considerations which the lay public may not always fully appreciate. Social-psychological surveys and the other methods described in this section have proven effective in identifying and measuring people's concerns, preferences and intentions and at uncovering underlying assumptions, knowledge, beliefs and feelings allowing conflicts between various publics and between

public preferences and ecological science to be better articulated and addressed by decision makers.

Many social-psychological methods were primarily developed and are frequently used for the purpose of determining characteristics and traits of people. For example, differences in the patterns of responses across a standard set of items in a questionnaire are used to ascertain and measure the intelligence, personality, religious beliefs, political leanings, and world views of individuals and groups. By shifting the focus from differences between persons (across items) to differences between items (across persons) the same basic instruments can be used to ascertain beliefs, attitudes, preferences, and behavioral intentions for a wide range of objects, events, states of the world and personal or social circumstances. For example, marketing surveys analyze responses to alternative goods and services to ascertain consumer preferences and purchasing intentions. Political polls collect responses to candidates and policy initiatives before elections to determine preferences and voting intentions. Social-psychological methods for assessing the value of ecosystems and ecosystem services focus on changes in environmental conditions, the associated social consequences of those changes, and on alternative means of achieving (or preventing) such changes. These assessments may also include measures of persons/groups, however, as when expressed ecosystems values are found to differ consistently between people with different beliefs, political leanings, or world views.

While public opinion is sometimes directly used to make policy decisions, as in elections and referenda, social-psychological assessment methods more typically are explicitly intended for decision support. Thus there has been little emphasis on mapping all expressed concerns and preferences onto a single metric scale (as in economic cost-benefit analysis methods, for example). More often separate indices are developed on several different value dimensions (e.g., aesthetic, ethical, utilitarian, personal, civic) for each policy alternative. Where alternative policies/outcomes fall on these psychological value scales may be seen as a precursor of resolving tradeoffs among competing values, for making choices between or for actually or hypothetically making monetary payments for them. Consistencies and conflicts between coherent sub-sets of the public (e.g., urban-rural, young-old, conservative-liberal, local-regional-national) in the relative orderings of alternatives on one or more value dimensions may also be estimated and reported. Differences among value

dimensions or between public groups are not typically resolved through aggregation or other calculation devices within the assessment process. Instead, such resolution is deferred to subsequent stages in the decision making process, where information integration, deliberation and negotiation may be applied in more or less formal procedures.

The social-psychological approach to assessing the value of ecosystems and ecosystem services enlists a number of different quantitative and qualitative methods. Formal surveys and questionnaires are the most frequently used and the most thoroughly researched set of quantitative methods. Surveys typically rely on standardized verbal descriptions of alternative objects/states (e.g., alternative environmental conditions or policies), with respondents recording explicit choices, rankings or ratings that can be quantitatively analyzed. Focus groups and individual narrative interview methods place more emphasis on qualitative analyses and have historically been used primarily in preliminary studies to support the design and pre-testing of quantitative surveys. These methods are less thoroughly researched than quantitative surveys, but they are being used with increasing frequency as stand-alone assessment methods. These qualitative methods typically employ less restrictive representations of options, are frequently directed at specific local cases that are familiar to respondents, and collect open narrative responses that are subjected to more or less rigorous qualitative analyses. In addition to these more established methods, some emerging methods base assessments on more direct observations of behaviors in the environments at issue. Behavioral observation and behavior trace methods have been developed and evaluated, especially in the context of the assessment of recreation and tourism values. Computer simulation ("virtual reality") and interactive game methods are also being developed, but have mostly been applied in research settings. These emerging methods are not yet sufficiently proven for immediate application in EPA policy-making contexts, but they do show considerable promise in areas where more traditional methods have been found lacking. They will only be briefly described in this section and are offered primarily as potential targets for future research and development.

3.2. Brief Description of the Methods

3.2.1 Surveys

Surveys encompass a broad range of methods for systematically asking people questions and recording and analyzing their answers (Schaeffer and Presser 2003).

Questions may assess knowledge, beliefs, desires and/or behavioral intentions about a virtually unlimited range of objects, processes, or states of the person, society or the world. The most popular survey formats have involved face-to-face, mail or telephone contacts with individually sampled respondents. Multiple questions/issues are presented and responses are typically reported as choices (among two or more options), rankings, or ratings (Dillman 1991; Krosnick 1999). Open-ended response formats are less often used, and pose special problems for quantitative analysis.

Social-psychological surveys have been extensively used to assess preferences, liking, acceptability and importance of presented policies, actions, outcomes and/or the expected personal or social consequences thereof. Multiple value dimensions (e.g., utilitarian, aesthetic, ethical) may be addressed within and between different surveys, and surveys may specify individual/personal, household/family or social/civic constituencies. The indices produced by application of appropriate quantitative analyses of recorded responses usually claim to be only ordinal (ranks) or roughly interval scale measures, providing relative measures of differences in the assessed values among the offered alternatives. Thus, generalization of obtained value measures (e.g., "values transfer") beyond the objects specifically assessed within the survey must be approached with caution. Moreover, expressed preferences or other value judgments are assumed to be at least in part created in the context of the survey (Schaeffer and Presser 2003).

Surveys have become ubiquitous in modern society, with uses ranging from assessments of diners' satisfaction with the service at a restaurant to citizens' support for major national policies (Dillman 2003). Surveys are now frequently directed by computer programs that can select and order questions individually for each respondent, sometimes based on responses to prior questions. Increasingly surveys are fully implemented by computer, allowing the respondent to control (with more or less restriction) the pace of questions and to record their responses directly into a computer database by key presses, clicks or voice commands (Tourangeau 2004). Internet-based methods offer extended possibilities for contacting respondents, presenting questions, and recording responses and their use is increasing. However, web surveys still have a number of hurdles to cross before being fully accepted in survey research (Couper 2001; Tourangeau 2004).

Variations on survey research methods that may be especially important to assessments of ecosystems and services include perceptual and conjoint survey methods. In perceptual surveys assessment targets (e.g., existing environmental conditions and/or projected policy outcomes) are represented by photographs, videos, computer visualizations, audio recordings, or even chemical samples representing different smells. As for verbal surveys, responses are typically choices, rankings or ratings of the offered alternatives. Perceptual surveys may be seen as extensions of traditional psychophysical research methods that have long been applied to assess qualities and preferences for foods and other products and environmental conditions that are difficult or impossible to describe effectively with words (Daniel 1990). Relevant examples include assessments of the visual aesthetic effects of alternative forest management policies in the northwestern US (Ribe et al. 2002, Ribe 2006), of in-stream flow levels on scenic and recreational values (e.g., Heatherington et al. 1993), of visibility-reducing air pollution on visitor experience in National Parks (e.g., Malm et al.1981), and assessment of the annoyance produced by aircraft over-flight noise in the Grand Canyon (Mace et al.1999). Perceptual surveys commonly also include more traditional verbal survey questions.

Surveys have most often presented the attributes of assessment targets separately. For example, a survey to assess the effects of a proposed environmental policy might present separate questions to determine respondent's judgments about effects on air quality, water quality and local employment Conjoint survey methods (e.g., Adamowicz et al.1998; Boxall et al. 1996) instead present options as multidimensional composites or scenarios presenting integrated combinations of different attributes (e.g., different levels of air quality, water quality and local employment, as illustrated in the accompanying box). Combinations generally reflect the actual or projected correlations among the combined attributes (e.g., air and water quality may be positively correlated, and both might be negatively correlated with local employment opportunities). In some cases, hypothetical combinations of attributes may be included, as specified by an experimental design to allow estimates of the separate and interacting effects of component attributes (Louviere 1988). Multiple regression analyses are used to estimate the relative contributions of individual components (attributes) to the expressed preferences (or other judgments) for the conjoint alternatives.

Conjoint survey methods can provide relatively direct estimates of the value tradeoffs people make when choosing among outcomes composed of multiple attributes that naturally covary and whose values potentially compete. When at least one of the attributes that forms the conjoint alternatives is (or can be) valued in monetary terms, the regression equation based on expressed preferences among the conjoint alternatives can be translated so that coefficients for all attributes are expressed as monetary values (e.g., Chattopadhyay et al. 2006, and see the discussion under economic assessment methods in this report).

Text Box 9: What Are Conjoint Surveys of Attitudes?

Conjoint methods may be especially well-suited for gauging public preferences across sets of complex multi-dimensional alternatives, such as alternative EPA regulations or management options for ecosystems/services protection. Respondents can be required to choose among (or rank or rate) compound alternatives that present specific packages of desired and less-desired attributes. For example, a policy that produces cleaner air and water in a region, but constrains employment opportunities in local communities might be pitted against alternatives that allow various levels of degradation in air and water quality, coupled with different levels of expanded employment opportunities. A simplified example of alternatives that might be presented to a respondent in a conjoint survey might be:

Which option do you think would be the best policy for public agencies in your area

Policy A: Resulting in a 10% improvement (from current conditions) in air quality, a 15% improvement in water quality, and a 15% decrease in local employment opportunities;

or

Policy B: Resulting in a 5% improvement (from current conditions) in air quality, a 10% improvement in water quality, and a 10% decrease in local employment opportunities.

Choices (or rankings or ratings) among a carefully constructed array of such alternatives can provide quantitative measures of relative public preferences for each policy option compared, as well as provide estimates of the contributions of each individual component or attribute to the conjoint preferences expressed. Following the simple example above, preferences for conjoint options might be represented by

$$\text{Preference for option } j = w1(WQj) + w2(AQj) + w3(Jobsj),$$

where option j is a particular policy that produces specific changes in the levels of water and air quality (WQj and AQj) and jobs (Jobsj). The relative contribution of each component/attribute is estimated by the derived coefficients (the wi) in the multiple regression equation for preferences among conjoint alternatives. Once determined, the regression equation can also be used to estimate preferences for new policy alternatives

(based on their respective projected measures of water and air quality and jobs), so long as those options fit within the range of the attributes assessed and the constraints imposed by the context of the survey in which the policy options were offered and judged. Optimization or less formal heuristics may be applied to create additional policy options for consideration and/or for direct evaluation in subsequent conjoint surveys.

3.2.2 Focus Groups

Focus group methods engage small groups of relevant stakeholders in facilitated discussion and deliberation on selected/focused topics relevant to the assessment of the effects of a policy, or alternative policies, outcomes and/or consequences. Typically experts and/or trained facilitators present the context, motivation and goals for the group and open-ended narratives are collected from the participants, usually in the context of discussion and deliberation with other members of the group and the experts/facilitators. Collected narratives are subjected to qualitative analyses to identify and possibly to ascertain levels of consensus on relevant issues, perspectives and positions represented by the participants. Reports of focus group results typically include numerous quotations of collected comments, along with the investigators' interpretations of the implications for the problems/policies/outcomes being addressed (e.g., Winter and Fried 2000). Less often collected narratives are subjected to more rigorous analyses based on formal logic models or discourse analysis systems (Abell 2004; Bennett and Elman 2006).

Relative to formal surveys, focus groups use small numbers of respondents and do not typically attempt formal probability sampling to select participants. Emphasis is instead on assuring that at least one representative from the full range of interests and perspectives relevant to the policies or outcomes at issue are included. The goal of a focus group is rarely value assessment per se, but a full discovery and articulation of all of the values that are relevant, and exploration of agreements and conflicts among the stakeholder constituencies represented by participants. Thus, focus groups are often employed early in policy and decision making, including the identification of the problems to be addressed and the formulation of alternative policies to address those problems. It is common for focus groups to be used in the process of designing and pre-testing more formal surveys (e.g., Shields et al. 2002).

3.2.3 Individual Narratives

The individual narrative method is distinguished from the focus group method in that individual respondents are contacted and participate alone, without interaction or discussion with experts, facilitators or other respondents. Included in this category are various ethnographic methods and the mental model methods described elsewhere in this report. Individuals nominally representing possible stakeholder perspectives are contacted and asked to comment on relatively broadly defined topics with relatively little direction from the interviewer/assessor (e.g., Brandenburg & Carroll 1995). Respondents are not typically selected by a random, probability sampling process. Instead, particular individuals are specifically targeted because of their known or assumed nominal group membership or personal relationship to the problem/policy/outcome at issue. The sample may be extended by having prior respondents refer others, as in the "snowball" technique. The number of individuals to be included is quite variable, and in a relatively few cases has been determined by some formal process based on a rolling analysis of collected narratives (e.g., using a criterion of diminishing new perspectives/positions being discovered). Collected narratives are subjected to more or less rigorous qualitative analyses, essentially similar to the analysis of focus group responses) to explore and articulate the breadth and depth of expressed understandings and concerns relevant to the assessment target.

3.2.4 Emerging Methods.

The social-psychological methods described in this section are relatively new and untested, but do show increasing potential for addressing value assessment issues not well handled by other methods. These methods are characterized by their attempt to more directly observe responses to policies, outcomes and consequences in situ, avoiding problems of relying on hypothetical responses to described hypothetical conditions. Moreover, observed environmental behavior is often not consistent with what people say they would do in the specified circumstances (or even with what they say they did) and people are often incorrect at identifying, or are unaware of the environmental factors that affect their behavior (Cole and Daniel 2004). In the context of ecosystems and services, behavioral observation methods monitor the activities of people in a particular environmental context and observe changes in behavior as relevant conditions change over time within a site or over sites with differing characteristics. Behavior trace methods are based on indirect evidence of people's

behavior in specific environmental contexts. For example, the number of visitors to recreation sites might be estimated by counting the number of autos parked at access points, by the number of passers-by recorded by automated trail counters, by the number of fire rings in dispersed camping areas or by the amount of trampling and disturbance of vegetation along trails and at destination points. Direct observations or traces of visitors' activities can be correlated geographically with relevant environmental/ecological conditions or monitored over time as changes in conditions occur at the same sites, revealing the effects of these changes on environmental preferences and reactions (e.g., Gimblett et al. 2001; Wang et al. 2001).

Interactive environmental simulation systems (sometimes approaching "virtual reality") provide means to overcome some of the limitations and difficulties of conducting direct behavioral observations or interpreting behavior traces. Direct observation methods are necessarily limited to existing conditions and are potentially confounded by uncontrolled or unrecognized irrelevant variables. Most policy decisions hinge on people's responses to specific changes to not-yet-existing, projected environmental conditions. Rapidly advancing computer technology has enabled effective and economical simulation of complex dynamic environments at high levels of realism (e.g., Bishop and Rohrmann 2003). The emphasis has been on visual presentations (now including high resolutions over full spherical view angles), but the technology can readily include auditory features and in some systems tactile, proprioceptive, olfactory, vestibular and other senses can also be effectively simulated to achieve very compelling, emersive environmental experiences. Moreover, expanding response options, ranging from the computer mouse to video-game controllers to gloves to full-body movement enable increasingly natural interactions with simulated environments. In the context of assessing the effects of changes in ecosystems and services, interactive computer simulation systems offer the opportunity to conduct virtual in situ experiments to determine how persons respond to specific investigator-controlled changes in environmental conditions. Thus the effects of manipulated conditions on environmental preferences and other reactions can be revealed in a context closely approximating "real world" circumstances.

In many regards interactive simulation systems come full circle back to the hypothetical responses to hypothetical environments that have challenged the validity of

verbal survey methods. However, as environmental simulations come into closer perceptual correspondence with actual environments and the available responses to simulated environments more closely approximate natural in situ responses, confidence in the validity of assessments based on these systems is increased. As for all other assessment methods, the gold standard is the empirical correspondence with the behavior of people in the actual environment of interest, if and when that environment is achieved.

Interactive computer simulation systems may be viewed as games, in which human respondents attempt to (virtually) navigate through and perhaps alter (virtual) environments to accomplish desired goals. There may be no particular outcome that can be defined as "winning" such a game, but the behavior of the player and the outcome on which s/he settles can reveal the values that motivate and guide the player's responses. Environmental games can be informative in this regard, even if they are played in substantially less than virtual environments. Indeed, more limited and/or more abstract games may have important advantages in some circumstances. For example, it may not be possible to project the explicit and detailed outcomes of a proposed policy that are required for a realistic environmental simulation, and the specific implications of particular responses to changing environmental conditions may not be known. In many situations only changes in some particular ecological component may be known and relevant (e.g., a reduction in a particular contaminant or an increase in survival rates of a particular wildlife or plant species). Still, a game-like context may be an effective and engaging way to communicate with public audiences about what outcomes they would prefer, and what policies are required to achieve those outcomes. A major advantage of games over surveys, for example, is the opportunity for respondents to learn through experience about how the ecosystem of interest responds to various policies or policy aspects and to progressively modify their expressed policy preferences to achieve some acceptable balance among desired and undesired outcomes.

3.3. Relation of Methods to the CVPESS Expanded and Integrated Assessment Framework

Survey methods have useful roles to play throughout the valuation process envisioned by C-VPESS. Surveys could contribute to initial problem formulation by identifying ecological services and impacts that most concern citizens and/or identified stakeholders, as well as by uncovering assumptions, beliefs and values that underlie that concern.

Importantly, similarities and differences in concerns among different segments of the public can also be identified and articulated. Once relevant ecological endpoints have been identified surveys could be very useful for determining the personal and social consequences of policy outcomes, and for exploring public understanding of the links between chains of ecological effects and the policy options under consideration (Box 2). Given a set of potential policy options, with their respective ecological endpoints (from Step 3), surveys could be used to assess relative public preferences (and/or other judgments, such as importance or acceptability) for those options (Step 4). Quantitative indices of public/stakeholder preferences (or judgments of importance or acceptability) from surveys could be combined with bio-ecological and economic/monetary measures of the value of the same alternatives to provide cross validation of all measures, or to identify possible limitations of either set of measures. Surveys may be especially useful when the values at issue are difficult to express or conceive in monetary terms or where monetary expressions/valuations are viewed as ethically inappropriate. In those cases social-psychological surveys could provide quantitative measures of public preferences among the policy alternatives or ecological endpoints that are under consideration, improving the basis for Agency decision making.

Surveys could make an additional contribution after Step 4 in the C-VPESS model. The values of ecosystems/services coming out of Step 4 must inevitably be represented by multiple economic/monetary, bio-ecological and social-psychological indicators. EPA administrators can be left with the difficult task of integrating these diverse and potentially conflicting measures, along with legal, budgetary and other constraints to make and rationalize policy decisions. Properly structured surveys, perhaps including material to inform respondents about relevant ecological and social effects and other considerations affecting the policy/decision at issue, could effectively involve citizen stakeholders in this value integration and tradeoff process, providing an additional relevant input to the policy decision, and adding to the political validity and social acceptability of the final action.

Focus groups and individual narrative methods would be most appropriate and most useful at the earliest and latest stages of the decision making process. While focus groups and individual interview methods do not generally provide quantitative assessments for alternative policies or outcomes, they can make important contributions to improving the

design, development and pre-testing of more formal surveys that can provide reliable and valid quantitative assessments of public concerns and values. Genuine probing interactions with individuals or groups representing key stakeholders and including divergent views and concerns should be a central part of problem definition and identification of significant ecological and associated social effects components of the process. Such interactions with key stakeholders and with citizens could also inform the values integration and negotiation in the final decision process and guide and pre-test the communication of that decision.

Status of the methods. Social-psychological surveys are the longest and most frequently used methods for determining public beliefs, concerns and preferences. Surveys have been and continue to be used effectively by all levels of government to ascertain citizen desires, concerns and preferences, by commercial marketers to determine the attractiveness of a wide array of goods and services, and by social and political scientists to measure and monitor shifting values and desires in the electorate. Economists have lately adapted survey methods to develop expressed preference methods for estimating monetary values for non-market goods and services, and surveys are often relied upon to collect the data needed to exercise other economic valuation efforts. Environmental management agencies have made extensive use of surveys, either directly or indirectly, in setting policy and in making and monitoring the effects of management decisions.

For a recent example of the application of survey methods in environmental management, the USDA Forest Service conducted a nationwide telephone survey to inform the Forest Service Strategic Plan, 2000 Revision, as required by the Government Performance and Results Act (Shields et al. 2002). By the authors' description, "approximately 7000 randomly selected members of the American public were asked about their *values* with respect to public lands, *objectives* for the management of forests and grasslands, *beliefs* about the role the USDA Forest Service should play in fulfilling those *objectives*, and *attitudes* about the job the USDA Forest Service has been doing in fulfilling their *objectives*" (p 1). This survey provided useful information about public values and concerns relevant to Forest Service management mandates, as well as quantitative measures of the relative importance to the public of particular policies (e.g., roadless areas, wilderness, timber harvesting, recreation opportunities, ecosystem health). Results were reported collectively and separately for different regions of the country, different demographic groups

and for groups evidencing in one section of the survey different levels of familiarity with the Forest Service and its management mandates.

The largest barriers to greater use of survey methods in the EPA are institutional. First, while the EPA seems to have embraced economic surveys (e.g., CVM, or at least "transfers" from prior CVM surveys) as a valuation method, there is a noticeable reluctance to use the larger class of systematic surveys, relative to the practices of other federal agencies with similar environmental protection mandates and valuation needs. This predisposition may in part be due to specific legal requirements for formal monetary benefit-cost analyses (which also apply to other agencies), but none of the currently applicable laws preclude using a fuller range of survey methods, and the most prominent laws and guides explicitly urge a broadly based evaluation effort not limited to monetary measures. Aside from this agency-level barrier, survey methods in general are discouraged by federal rules implementing the Paperwork Reduction Act. Over the past several decades it has been very difficult for federal agencies to attain required clearances (e.g., from the OMB) for surveying the public in a manner and in a time frame that effectively addresses policy evaluation needs. This institutional barrier is formidable, and the proliferation of surveys and pseudo-surveys has dampened citizen's willingness to participate, but many significant surveys continue to be conducted by a number of government agencies.

It is not clear the extent to which focus groups or individual interviews are systematically used in EPA policy making, nor do the OMB and other guidelines clearly specify the criteria for using these methods. Focus groups are widely used in marketing and political polling contexts and the US Forest Service national survey by Shields et al (2002) described above reported that "over 80 focus groups conducted around the continental United States" (p. 1) were used in the design and development of the survey, as well as to support the interpretations and conclusions from the survey. "Public meetings" and on-site demonstrations are frequently cited as playing a public involvement role in EPA policy decisions, and a formal "Multi-Stakeholder Group" was assembled and used in the Avtex Fibers Superfund Site decision and implementation process (cite), but it is not clear whether any of these activities can be construed as using a focus group, nor is it clear how often such methods have been used to systematically compare alternative policies/actions.

The use of focus groups would seem to be completely consistent with previous advice of the EPA Science Advisory Board (US EPA 2001) recommending increased use of "stakeholder processes" in Agency decision making. Stakeholder processes were defined as "…group processes in which the participants include non-expert and semi-expert citizens, and/or representatives of environmental non-governmental organizations, corporations and other private parties in which the group is asked to work together to: define or frame a problem; develop feedback in order to better inform decisionmakers about proposed alternative courses of action; develop and elaborate a range of options and/or criteria for good decision-making which a decision-maker might employ; or, either explicitly or implicitly, actually make environmental decisions." (p 8) Still, the term "focus group" was not used anywhere in this document. While no specific evidence has been found either way, it seems reasonable to assume that individual narrative interviews have not been important components of EPA decision making processes. Certainly the qualitative nature of the information provided by both focus groups and individual interviews, and the general disinterest in representative sampling makes them poor candidates for formal policy evaluation exercises, but that does not preclude their having a role in earlier stages of the decision making process as envisioned by the C-VPESS.

Limitations. Surveys have proven very effective for determining public knowledge, beliefs, attitudes and intentions. However, especially in the context of the complex processes of selecting alternative policies and actions to protect ecosystems and services it is important to recognize that the responding public may not have a sufficient basis for the opinions and preferences offered in a general population survey. First, limitations on length and complexity of content (especially for telephone surveys) make it unlikely that the full complexity, including uncertainties of policies and their outcomes can be effectively communicated to respondents within the survey. Second, the general public is unlikely to have the breadth and depth of ecological knowledge that is often required to understand and evaluate a given policy, its bio-physical outcomes or the implications of outcomes for the respondent or for society more generally. Finally, even when the respondent fully understands these aspects of a proposed policy he/she may still be uncertain (or incorrect in his/her projection) regarding how well (or badly) the respondent will feel about the outcomes/implications when they are actually encountered.

Approaches to the problems of insufficient communication of the options within the survey, and insufficient relevant prior knowledge/experience of the respondents include increasing considerably the information provided within the survey, requiring respondents to complete some workshop, training exercises or participate in guided visits to actual demonstration sites before responding, or shifting to a much more intensive interactive "deliberative group" format (see section X in this report) that includes input of relevant experts and extended discussions among respondents. While these approaches may be successful in informing and educating respondents, they may also invalidate generalizations to the population from which respondents were originally sampled—i.e., the responses of informed, educated respondents can only represent the population of similarly informed and educated publics. This presents something of a dilemma for survey research (and for public involvement more generally) in the context of public agency policy and decision making. If public beliefs, preferences and intentions regarding a proposed policy are important—as in predicting the level of public acceptance and support the policy will encounter, what objections might be raised against it, and how likely are desired responses (e.g., compliance with regulations, behavior changes needed to affect the policy, voting for political candidates that will help to sustain the policy, etc)—then, given that the preponderance of the public will be less well-informed and less educated about the policy and its implications than the survey respondents, the results of the enhanced survey may not be very useful. On the other horn, the opinions and preferences of uninformed and uneducated (in the specific issues relevant to the policy) respondents would seem a poor basis for deciding ecological protection policies.

Perhaps it would be better to leave the public out of the process altogether and have experts (e.g., ecologists, biologists, toxicologists) determine ecosystem and ecosystem service protection policies. However, it is not always clear what the necessary normative principles are for such decisions, or who should decide among competing principles. In some cases, there may be high levels of agreement among experts about the bio-physical outcomes of proposed policies, and even about the implications of those outcomes for individual and social well-being. However, when expert judgments on these matters are incongruent with the beliefs, preferences or intentions of the public, there is a problem. The involvement of citizens in decisions about their future environments and what would best serve their individual and collective well-being seems a basic tenant of democratic societies.

This does not require that complex environmental policies and actions be determined by a public opinion poll (and surveys consistently show that the public does not expect or want this), but certainly public opinion always matters to public policy. In this context, surveys of the less intensive type can be very helpful in identifying where publics and experts disagree and why, and what information, deliberations or other interventions might be needed to achieve better congruence of expert and public opinion.

Research has shown that survey respondents faced with questions that imply the need for answers will usually respond as best they can under the circumstances. This often means that respondents will "grasp at straws" provided by the survey itself (e.g., who is conducting the survey, the apparent goals/intentions, the content of other/prior questions, what answers they have already recorded) to "construct" their answers (and perhaps the presumed underlying values) on the spot. For this and other reasons it is very important that surveys are carefully designed to present an unbiased representation of policy options in a context that is as close as possible to the actual contexts in which the policies/outcomes/implications being assessed will be encountered in the "real world." The ultimate test of the validity of a survey is how closely the responses collected in the survey correlate with responses made by the population of interest when the policy is actually implemented—e.g., the policy most favored in the survey should be the policy most favored in the "real world."

The technical issues that have been of the greatest concern to users of survey information, to quality control agents (e.g., OMB) and to survey researchers have been associated with the sampling of respondents. The results of a survey are typically intended to be generalized to some specified population (e.g., adult citizens of the US) that includes many members that will not be included in the sample of individuals who actually respond to the survey (i.e., the respondents). The integrity of generalizations to the population of interest is assured if the respondents are a formal "representative sample" of the population (i.e., classically, every member of the population has an equal chance of being a respondent, or in contemporary survey research terminology, a "probability sample" is used). Because every element of the population of interest may not be available or even in principle known, the first step in attaining a probability sample is to establish a sampling frame, or an accessible list from which a smaller number of individuals (or households) will be randomly selected and contacted. At this stage errors or "biases" may enter to the extent that the

sampling frame does not properly represent the entire population of interest (e.g., if the sampling frame is telephone numbers listed in directories, otherwise qualified individuals/households with unlisted numbers will not have an equal chance to participate in the survey).

Because not all individuals/households selected from the sampling frame will successfully be contacted, and some who are contacted will choose not to actually participate in the survey, there is an opportunity for "non-response error" to degrade or bias the generalization of results to the population represented by the sampling frame. In spite of the attention paid to non-response errors in surveys, research shows that the biasing effects of non-response are typically negligible. That is, the results of surveys with low response rates (e.g., on the order of 20%) usually provide a sufficient basis for valid inferences about the population represented by the sampling frame. In part this can be attributed to the fact that the reasons for non-participation are usually not systematically related to the subjects addressed in the survey. More difficult and potentially more potent errors are in survey design, including the crafting, selection and ordering of questions/items to be included in the survey, the form of the response options offered (e.g., the type of ratings scales) and uncontrolled events that occur during the time of survey implementation (see Krosnick 1999 and the appendix to this report).

Social-psychological surveys do not meet the requirements of conventional economic cost-benefit or cost-effectiveness analyses because they do not achieve a unidimensional, transituational measure of value. That is, the scale values computed for the ecosystem and service options addressed in a survey can not be directly compared to values for extra-survey options, or to values and costs in other domains of the respondents' lives. However, given the identification of a feasible set of alternative regulatory/protection actions, these survey methods would be appropriate for assessing public preferences among offered sets of policy/outcome options, and for estimating relationships among the multiple component attributes (costs and benefits, biological and social effects, means and ends) of those actions. Properly designed conjoint methods may be especially well-suited for gauging public preferences across sets of complex multi-dimensional alternatives, such as alternative EPA regulations for ecosystems/services protection.

In practical use, the human resources required to implement surveys range from a sufficient cadre of technically competent survey designers and analysts to temporary hourly wage employees to perform the mailing, phoning or interviewing tasks. Material needs may be very low ("paper and pencils") or quite high, as when sophisticated computer simulations/visualizations or interactive response formats are employed. Face-to-face surveys, where trained interviewers are required and contact costs may be high, are generally the most expensive, but costs for mail, telephone and/or computer resources can be significant in large surveys using those formats. All of these costs are usually quite low relative to the physical, biological and/or ecological science and field study required to create adequate projections/representations of value-relevant means and outcomes for a suitable range of alternative regulatory or protection actions. In many ways, the quality of evaluations of ecosystem and ecosystem service protections most depends upon the quality of the relevant projections and specifications of ecological endpoints and their social consequences. In some cases considerable resources may have to be devoted to translating targeted ecological outcomes into understandable representations of socially relevant effects. Thus, once these essential factors have been accomplished, the cost of the actual survey can be comparatively quite small.

Focus groups and individual interviews can have important and useful roles to play in Agency policy and decision making. However, their emphasis on qualitative analyses and their typical disregard for representative sampling can make them less useful for systematic evaluations or comparisons of alternative policies and outcomes. These methods can very useful and important for designing and pre-testing more formal surveys that do provide quantitative assessments of values for alternative policies and outcomes. Qualitative methods may also contribute to the design of more effective communications and rationalizations of Agency decisions to stakeholders and to the general public.

Treatment of Uncertainty. There are two broad levels of uncertainty in any evaluation of changes in ecosystems and ecosystem services. At the bio-physical level any characterization of current (or past) ecological conditions will have numerous interrelated uncertainties, and these uncertainties will be magnified and added to by the need to project future conditions, with or without some postulated management action. At the social level existing and projected ecological conditions and their socially relevant consequences must be

represented to people so that they can express their preferences for current and/or alternative future conditions, providing the basis for analysts to derive valid measures of the value of the targeted changes. Within the social level, survey methods specifically must address sampling errors (e.g., representative sampling, non-response), specification errors (e.g., adequate description/representation of alternatives, clear and understandable response system) and the effects of a variety of contextual and external factors that may affect (bias) participant responses. Methods for reducing and quantifying the magnitude of most of these sources of uncertainty and error in surveys are part of the well-documented technology and the accumulated lore of survey research (e.g., Dillman 1991, Krosnick 1999, Tourtangau 2004, and the appendix attached to this report).

While the currently available methods for dealing with uncertainty may be sufficient for some simple evaluation problems, the valuation of changes in ecosystems and ecosystem services raises issues not well addressed by any existing methods. For example, at the bio-physical level it is extremely difficult or completely unclear how to calculate the uncertainty (error) in the projection of even a single outcome (endpoint) from a complex ecological system composed of multiple interacting variables that may be separately non-linear and collectively subject to the influence of unpredictable external stochastic events. Modeling methods, such as Monte Carlo and other sensitivity analyses may be used to estimate the distribution of possible outcomes (or at least best-case, worst-case extremes) for a single endpoint, but even this approach becomes unwieldy when the outcomes relevant to the value assessment are themselves composed of multiple interrelated variables. While highly trained and experienced experts may find ways to calculate a relevant measure of uncertainty for some complex ecological outcomes, it is problematic how to meaningfully communicate this level of uncertainty to concerned lay citizens in a survey. Yet such communication can be crucial, as often the level of uncertainty in outcomes is a key factor affecting preferences for the alternative policies under consideration.

Accepted methods are available and are commonly used for calculating confidence intervals or complete probability distributions for individual survey responses over respondents (e.g., the importance ratings assigned to a particular item). The internal reliability and cohesiveness of survey responses can be calculated per individual respondent, but more often the focus is on the mean response of homogeneous groups of respondents.

Multiple items are frequently combined, as by cluster or factor analysis, into latent variables (factors) implied by the inter-correlations among individual-item responses, and there are several conventional statistical indices of the internal consistency and coherence of those derived factors. More complete analyses calculate and quantitatively assess the internal consistency and distinctiveness of latent variables, based on the patterns of responses across the multiple respondents, as well as classifying sub-groups of respondents, based on patterns of individual's responses to the multiple items in the survey.

The detailed results of a complex attitude survey are unlikely to be fully appreciated by anyone without relevant training and experience. On the other hand, results can be, and routinely are simplified for communication to lay audiences. Most people would find reports such as "alternative A was preferred over all others offered in the survey by 75% of respondents" to be clear and intuitively understandable. A table or graph showing mean preference ratings on a 10-point scale for all alternatives evaluated would be clear to many members of the public, as well as to experts from other scientific and managerial disciplines that are involved in EPA rule and decision making. Some of the uncertainty associated with these indices (e.g., the sampling error) could be displayed by conventional confidence intervals or error bars. The potential effects of more complex sources of uncertainty might be revealed by bracketing mean estimates for each alternative assessed with 25th and 75th percentile estimates derived from sensitivity analyses exercised over the entire biological-social evaluation system. The most sophisticated communication devices might be based on interactive game systems, where the audience is allowed to alter input variables and assumptions about functional relations and stochastic events and observe and learn for themselves how these changes affect evaluation outcomes.

<u>Research needs</u> (still to be completed, but roughly addressing the following items)

- How can social-psychological surveys best be used in EPA policy and decision making, including how decision makers can and should use the relative quantitative (non-monetary) value indices provided?
- How can social-psychological value indices be used to cross-validate estimates of monetary values (e.g., from CBA) for alternative policies/outcomes?

- How can social-psychological, economic and bio-ecological evaluations of changes in ecosystems and ecosystems services for alternative policies/outcomes most effectively be integrated to support Agency policy and decision making?
- What productive roles can focus groups, individual interviews and other qualitative methods play in Agency policy and decision making?
- How might emerging methods (behavior observation, behavior trace, interactive computer simulations and games) be shaped to effectively contribute to Agency policy and decision making needs?

Key References

Abell, P. 2004. Narrative explanation: an alternative to variable-centered explanation? Annual Review of Sociology, 30:287–310

Adamowicz, W., Boxall, P., Wilhams, M., & Louviere, J. (1998). Stated preference approaches for measuring passive use values: Choice experiments and contingent valuation. American Journal of Agricultural Economics, 80, 64-7

Bennett, A. and Elman, C. 2006. Qualitative research: recent developments in case study methods. Annual Review of Political Science, 9:455–76

Bishop, I. D. & Rohrmann, B. (2003) Subjective responses to simulated and real environments: a comparison. Landscape and Urban Planning, 65: 261-267.

Brandenburg, A.M. & Carroll, M.S. (1995). Your place or mine? The effect of place creation on environmental values and landscape meanings. Society & Natural Resources 8(5): 381-398.

Chattopadhyay, S., Braden, J. B. and Patunru, A. 2005. Benefits of hazardous waste cleanup: new evidence from survey- and market-based property value approaches. Contemporary Economic Policy, 23, 357-375.

Cole, D. N. & Daniel, T. C. (2004) The science of visitor management in parks and protected areas: from verbal reports to simulation models. Journal for Nature Conservation, 11, 269-277.

Couper M. P. 2001. Web surveys: a review of issues and approaches. Public Opinion Quarterly, 64, 464–94

Daniel, T. C. Measuring the quality of the human environment: a psychophysical approach. American Psychologist, 1990, 45, 633-637.

Daniel, T.C. & Gimblett, H.R. (2000) Autonomous agents in the park: an introduction to the Grand Canyon River Trip Simulation Model. International Journal of Wilderness, 6: 39-43.

Dillman D. A. 1991. The design and administration of mail surveys. Annu. Rev. Sociol. 17:225–249.

Dillman D. A. Navigating the rapids of change: some observations on survey methodology in the early twenty-first century (2002) Public Opinion Quarterly Volume 66:473–494.

Dunlap, R. E., Van Liere, K. D., Mertig, A. G. and Jones, R. E. 2000. Measuring endorsement of the New Ecological Paradigm: a revised

NEP scale. Journal of Social Issues, 56, 425-442. Gimblett, H. R., Daniel, T. C., Cherry, S. & Meitner, M. J. (2001) The simulation and visualization of complex human-environment interactions. Landscape & Urban Planning, 54, 63-79.

Hetherington, J., Daniel, T.C. and Brown, T.C. (1993) Is motion more important than it sounds? The medium of presentation in environmental perception research. Journal of Environmental Psychology, 13, 283-291.

Kelly, J., P. W. Williams, A. Schieven and I. Dunn. 2006. Toward a Destination Visitor Attendance Estimation Model: Whistler, British Columbia, Canada. Journal of Travel Research, Vol. 44, 449-456

Krosnick JA. 1999. Survey research. Annual Review of Psychology, 50:537–67

Louviere, J.J., 1988. Analysing decision making: metric conjoint analysis. Sage University Papers Series in Quantitative Applications in the Social Sciences, N8 67 Newbury Park, CA: Sage.

Mace, B. L., Bell, P. A. and Loomis, R. J. 1999. Aesthetic, affective, and cognitive effects of noise on natural landscape assessment. Society & Natural Resources; Apr/May99, Vol. 12 Issue 3, p225, 18p, 5

Malm, W., Kelly, K., Molenar, J., & Daniel, T. C. Human perception of visual air quality: Uniform haze. Atmospheric Environment, 1981, 15(10/11), 1874-1890.

Ribe, R. G. 2006. Perceptions of forestry alternatives in the US Pacific Northwest: information effects and acceptability distribution analysis. Journal of Environmental Psychology, 26, 100-115.

Ribe, R.G., Armstrong, E.T., Gobster, P.H. (2002) Scenic vistas and the changing policy landscape: visualizing and testing the role of visual resources in ecosystem management. Landscape Journal, 21: 42–66.

Schaeffer, N. C. and Presser, S. 2003. The science of asking questions. Annual Review of Sociology, 29:65–88

Shields, D. J., Martin, I.M., Martin, W.E., Haefele M.A. (2002) Survey results of the American public's values, objectives, beliefs, and attitudes regarding forests and grasslands: A technical document supporting the 2000 USDA Forest Service RPA Assessment. General Technical Report, RMRS-GTR-95. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. 111 p.

Tourangeau, R. 2004. Survey research and societal change. Annual Review of Psychology, 55:775–801

U S EPA, Office of Policy, Economics and Innovation. 2003. Public Involvement Policy of the U.S. Environmental Protection Agency, EPA 233-B-03-002 http://www.epa.gov/policy2003/policy2003.pdf

US EPA Science Advisory Board (2001) Improved science-based environmental stakeholder processes: A commentary by the EPA Science Advisory Board. EPA-SAB-EC-COM-01-006

Wang, B. & Manning, R.E. 2001. Computer Simulation Modeling for Recreation Management: A Study on Carriage Road Use in Acadia National Park, Maine, USA. Environmental Management 23(2): 193–203.

Winter, G. & Fried, J.S. (2000) Homeowner perspectives on fire hazard, responsibility, and management strategies at the wildland-urban interface. Society & Natural Resources 13: 33-50.

3.4. Transfer of Socio-Psychological Information

Text to be inserted here

4. ECONOMIC METHODS

4.1. Overview

Brief Description of Methods: The economic concept of value is based on two fundamental premises of neoclassical welfare economics: that the purpose of economic activity is to increase the well-being of the individuals in the society and that individuals are the best judges of how well off they are in any given situation and what changes would enhance that well being. A number of economic methods have been developed for assessing changes in the well-being of the members of society, where well-being is defined as the individuals' preferences and their willingness to pay for gains or to accept compensation for losses.

Is this the place to insert something about Paul Slovic's view about preferences – as a controversial qualification to the usefulness of economic methods? Or is that going to be done somewhere else? – Rick F.

Substitutability is at the core of the economist's concept of value. This is because substitutability establishes trade-off ratios between pairs of goods that matter to people. The trade-offs that people make as they choose less of one good and substitute more of another good reveal the values that people place on these goods. If one of the goods has a monetary price for each unit that is consumed, then the revealed can be expressed in monetary terms. The money price of a market good is a special case of a trade-off ratio because the money spent to purchase one unit of one element of the bundle is a proxy for the quantities of one or more of the other elements in the bundle that had to be reduced to make the purchase.

A measure of the economic value of an environmental change based on substitutability can be expressed either as a compensating surplus (CS) or an equivalent surplus (ES). In principle, CS and ES measures can be defined in terms of any other good an individual is willing to substitute for the good being valued. In the following discussion, money is used as the numéraire in which trade-off ratios are expressed; but CS and ES could be measured in terms of any other good that mattered to the individual. The CS for an environmental improvement is the maximum sum of money the individual would be willing to pay rather than do without the

improvement. This sum is the amount of money that would make the individual indifferent between paying for and having the improvement and foregoing the improvement, while keeping the money to spend on other things. The CS for an improvement is also known as willingness to pay (WTP).

The ES for an improvement is the minimum sum of money the individual would require to voluntarily forego the improvement; it is the amount that would make a person indifferent between having the improvement and foregoing the improvement, while getting the extra money. In other words, it is the amount of money that would generate an increase in utility equivalent to that realized from the improvement in the environmental amenity. The ES for an improvement is also known as willingness to accept (WTA) compensation.

The key input for all of the economic methods is data on the choices that people make about the things that contribute to their economic well-being. These choices are made in several contexts. The first is choices about quantities demanded and supplied in markets at alternative prices. These choices generate demand and supply functions that can be estimated with the information on the amounts purchased at different prices using econometric methods. Changes in these demand and supply functions in response to changes in the levels of ecosystem services can be analyzed to obtain market-based estimates of the values of the changes in these services. Second, choices can involve the selection of quantities of nonmarket goods and services or responses to changes in the availability of nonmarket goods and services such as ecosystem services. Nonmarket revealed preference methods can be used to obtain estimates of the values of these changes. And third, choices made in response to hypothetical questions can be analyzed with one of the several stated preference methods for valuation.

In all cases, the outputs are monetary measures of economic value that allow a direct comparison of the benefits and costs of alternative policy choices in a common metric.

<u>Where do the Methods Fit in the CVPESS Integrated Framework</u>? The economic methods fit into step 4 of the integrated framework: Characterization of the Value of the Change in Monetary Terms.

Limitations: The logic underlying the definition of an individual's willingness to pay implies that the maximum amount a person could pay for anything is constrained by that person's ability to pay which is indicated by the individual's wealth. Thus the value estimates derived from economic valuation methods are conditional on the existing distribution of income and prices. And acceptance of these benefit estimates implies acceptance of the underlying distribution of wealth. One way to incorporate concern for equity in the distribution of well-being, with roots going back to Bergson (1938), is to weight the measures of economic value or welfare change for each individual by that person's relative degree of "deservingness"; that is, to employ a social welfare function that attaches a higher weight to benefits going to those judged to be more deserving because of some attribute such as their lower level of income. However, the search for a social welfare function has been unsuccessful. In practice, analysts typically use the value measures derived from the mean individual in the sample that is providing data for the valuation model in use. If value or willingness to pay is an increasing function of income, the analyst is implicitly underestimating the values of the highest income individuals and overestimating the values of the lowest income individuals. The result, in a crude qualitative sense at least, is equivalent to what would be obtained with a social welfare function.

Uncertainty: All of the available techniques for reflecting uncertainty are available for use with estimates of value based on economic methods.

Key References:

Bergson, A. 1938. A Reformulation of Certain Aspects of Welfare Economics. Quarterly Journal of Economics 52:310-334

Bockstael, Nancy E., and A. Myrick Freeman III. 2005. "Welfare Theory and Valuation," in Karl-Goran Maler and Jeffrey R. Vincent, eds., Handbook of Environmental Economics, Amsterdam: Elsevier.

Champ, Patricia A., Kevin J. Boyle, and Thomas C. Brown, eds.. 2003. A Primer on Nonmarket Valuation, Dordrecht: Kluwer Academic Publishers.

Freeman, A. M. III. 2003. The Measurement of Environmental and Resource Values: Theory and Methods. 2nd ed. Washington, DC: Resources for the Future

4.2. Market-Based Methods

Brief Description of Method: The market-based approaches to economic valuation are used to estimate the economic values of ecosystem services that are an input into the production of a good or service that can be bought and sold in a market at an observable price. Wetlands, for example, often serve as nurseries for fish species that are harvested for commercial markets. They are thus an input to commercial fishing and their services affect the supply and market price of harvested fish. The economic benefits of protecting wetlands can then be estimated by their contribution to the output of the commercial fishery. Similarly, when a river is used as a source of water for a municipal drinking water plant, the river's water quality directly contributes to the production of drinking water. A change in water quality can be assessed for its potential effect on the production and cost of drinking water.

When the final good is a commercially marketed good or service, these changes can be valued using market information about prices. For private goods and services purchased in competitive markets, the price of a good reflects the set of participants in that market valuation of an extra unit of that good or service. For small changes, market prices can be used as a measure of economic value of each unit of the goods involved. For larger changes, however, marginal willingness to pay (demand) and marginal cost (supply) are unlikely to remain constant requiring estimation of changes in consumer and producer surplus. [38]

Status as a Method: Market-based methods are based on well established economic principles and econometric practices (Boardman, et al., 2006, McConnell and Bockstael, 2005). They have been used for more than 30 years to evaluate a variety of economic policies including controls on prices and the benefits of deregulating markets (Hufbauer and Elliott, 1994, Winston, 1993). Applications to the valuation of ecosystem services include Barbier and Strand (1998) and Barbier, Strand, and Sathirathai (2002). EPA has used these methods to value ecosystem service benefits from air pollution control in the markets for agricultural products and for timber products (US EPA, 1999).

Limitations: Estimating both consumer and producer surplus requires the development of empirical models for the demand and supply relationships describing market outcomes. Depending on each application this can be difficult due to lack of data at the level of resolution required to describe how economic policies affect each of these relationships.

The majority of environmental policies do not directly impact the prices and quantities of goods and services traded in markets, so this method is only available in a limited subset of cases. Another limitation of this method is that market imperfections can confound the measurement of demand and supply and can distort the relationship between prices and the marginal value and marginal cost of providing a private good. As a result this distortion will carry over into any estimation of economic values based on market prices.

There are many non-environmental factors that can affect demand and supply relationships that are also important to assuring that policies are authentically reflected in the models. Seasonal variations in use or availability of goods and services related to environmental policies can affect prices, and this needs to be sorted out. The modeling and estimation of demand and supply functions can be complicated. Ultimately what can be learned about the influence of environmental or any other policy is limited by the available data. These limitations are best described as an identification problem – do we have sufficient information to identify the effects that are hypothesized to reflect how environmental policy influences market supply and demand?

Key References

Barbier, Edward B., and Ivar Strand. 1998. "Valuing Mangrove-Fishery Linkages," Environmental and Resource Economics, 12:151-166.

Barbier, Edward B., Ivar Strand, and Suthawan Sathirathai. 2002. "Do Open Access Conditions Affect the Valuation of an Externality? Estimating the Welfare Effects of Mangrove-Fishery Linkages in Thailand, Environmental and Resource Economics, 21:343-367.

Boardman, Anthony E., David H. Greenberg, Aidan R. Vining, and David L. Weimer. 2006. Cost-Benefit Analysis: Concepts and Practice, third edition Upper Saddle River, NJ: Prentice-Hall.

Hufbauer, Gary, and Kimberly Ann Elliott. 1994. Measuring the Costs of Protection in the US, Washington, DC : Institute for International Economics.

McConnell, Kenneth E., and Nancy E. Bockstael, 2005. "Welfare Theory and Valuation," in Karl-Goran Maler and Jeffrey R. Vincent, eds., Handbook of Environmental Economics, Amsterdam: Elsevier.

US EPA. 1999. The Benefits and Costs of the Clean Air Act 1999 to 2010, Washington, DC.

Winston, Clifford. 1993. "Economic Deregulation: Days of Reckoning for Microeconomists," Journal of Economic Literature.

4.3. Non-market Methods – Revealed Preference

When market prices and quantities are not directly affected by environmental policies, non-market valuation, using either revealed preference or stated preference, becomes necessary. Revealed preference methods look at people's behavior in markets that are related to ecological services to reveal underlying values. Because the revealed preference methods for measuring values use data on observed behavior, some theoretical framework must be developed to model this behavior and to relate the behavior to the desired monetary measures of value and welfare change. A key element in the theoretical framework is the model of the optimizing behavior of an economic agent (individual or firm) that relates the agent's choices to the relevant prices and constraints, including the level of ecological services being provided. If a behavioral relationship between observable choice vari¬ables and the ecosystem service can be specified and estimated, this relationship can be used to calculate the economic value of changes in these service flows.

The degree to which inferences about the value of a change in ecosystem services can be drawn from market observations, and the appropriate techniques to be used in drawing these inferences, both depend on the way in which the ecosystem service enters individual utility functions. The exploitation of possible relationships between environmental goods and private goods leads to several empirical techniques for estimating environmental and resource values. These techniques have the following characteristics: (1) they are consistent with the basic theory of demand and consumer preferences; (2) they provide a means for estimating the indirect utility function, the expenditure function, or the compensated demand function for the environmental service; and (3) they are practical in the sense of imposing realistic data and computational requirements.

This section covers three revealed preference methods: travel cost, hedonics, and averting or mitigating behavior models.

4.3.1 Travel cost

Brief description of the method: The travel cost method accepts as a maintained hypothesis that people have economic demand functions for the services of environmental resources that are associated with outdoor recreation. Lakes, rivers, forests, beaches, etc. are examples of the types of resources involved. The essence of the method is recognition that users pay an implicit price by giving up time and money to take trips to these areas for recreation. This recognition is important because most of the public facilities for recreation in the U.S. do not have market determined fees for that use. The cost of a visit to a site is the out-of-pockets costs of travel including any site admission fees, opportunity cost of travel time, and the opportunity cost of time on site.[39]

The values of ecosystem services are captured by the method to the extent they can be represented as factors that influence a person's demand (or site choice). For example, a measure of the availability of fish in a lake used for fishing would presumably influence a person's demand for a visit to the site for fishing. Sometimes the loss of an area (due to activities that eliminate its recreational value) is represented as "equivalent to" a price or travel cost change that is large enough to cause all existing users to no longer take trips to the site.

The travel cost demand literature in this area until about the middle 1990s estimated travel costs for the simple case of a new site or loss of site. To use the travel cost method for more sophisticated environmental policy choices, i.e. those that change the quality of recreational opportunities, analysts need to know how those quality attributes influence the demand function for recreational resource. In practice most economic models for recreation now use random utility models (RUM) which describe the decision process associated with each individual selecting which recreation site among a number of alternatives to visit. A RUM framework describes these choices as the result of a constrained optimization process; that is, selecting the site that yields the maximum level of utility (or well being) that is possible given a person's constraints. The result can be expressed as a function of travel costs, site characteristics including the level of ecosystem services, facilities to support specific activities (e.g. boat ramps, ski lifts etc), and users' attributes.

Status as method: It is based on well-established economic principles. There has been extensive use of this method in peer-reviewed literature, dating to 1947 when it was first proposed by Harold Hotelling. There is less experience with using the method to estimate tradeoffs for a wide range of attributes of recreation sites. Most of the experience until the middle 1990s was in evaluating changes in travel costs or the tradeoffs to obtain new sites or avoid losing existing ones. Assumptions are understood and documented. Meta analyses – Smith and Kaoru [1990],Walsh, Johnson and McKean [1992], Rosenberger and Loomis [2000], Johnston et al. [2003] and ,Johnston et.al.[2005]– have documented the performance of the model in different circumstances.

Measures of the economic value have been used in EPA's RIA analyses for regulations affecting recreation resources. A recent example is the Phase III component of the 316B rule. The rule seeks to reduce impingement and entrainment of fish and other organisms through power facilities' uptake of cooling water.

Strengths and Limitations: The primary data requirements are: data on people's usage of recreation sites; measures of individuals' values of time and time constraints; information that allows measures of the environmental attributes of the resources used for recreation to be linked to those resources; and information that describes the relationship between technical indexes of the attributes of recreation sites and measures that users can be expected to understand and know.

The analysis requires technical training in micro-economic modeling of demand and extensive experience with micro-econometrics to estimate recreation demand models. Less experience is required to use existing models to estimate economic values for changes in factors hypothesized to affect people's recreation behavior.

Uncertainties: One important source of uncertainty in the travel cost model is the value of recreationists' time as a component of the cost of a recreation trip. Randall has argued that for several reasons "travel cost is inherently unobservable" (1994, p. 88). The role of time in explaining recreation demand and in valuing recreation visits and sites raises some thorny issues for both the standard travel cost and RUM approaches of analysis. Clearly, time is an important variable in the analysis of recreation demand and value. But numerical estimates of demand and value require either that the numerical value of the shadow price of time be known or can be estimated from a model of the choices made

regarding the uses of time. A variety of models of choice and time are available in the literature. But as yet, different model structures yield quite different estimates of the shadow price of time. And there is no clear basis for preferring one model and its value over other models. Until these issues can be resolved, estimates of recreation values should be presented as conditional upon a specific value of the shadow price of time or a specific modeling approach regarding the role of time. And the uncertainty in the estimates that this implies should be acknowledged. For more on this issue, see Freeman (2003, Ch. 13).

Key References

P.A. Champ, K.J. Boyle and T.C. Brown, editors, A Primer on Non-Market Valuation (Dordrecht: Klimer Academic 2003).

A.M. Freeman III, The Measurement of Environmental and Resource Values, second edition (Washington, D.C. Resources for the Future 2003).

Haab, T.C. and K.E. McConnell, 2002, Valuing Environmental and Natural Resources, Cheltenham, UK: Edward Elgar.

Johnston, Robert J., Elena Y. Besedin, and Ryan F. Wardwell, 2003, "Modeling Relationships Between Use and Nonuse Values for Surface Water Quality: A Meta-Analysis," Water Resources Research, 39(12).

Johnston, Robert J., Matthew H. Ranson, Elena Y. Besedin, and Erik C. Helm, 2005, "What Determines Willingness to Pay per Fish? A Meta-Analysis of Recreational Fishing Values," under review at Marine Resource Economics.

D.J. Phaneuf and V.K. Smith. 2005. "Recreation Demand Models," in K. Mäler and J. Vincent, editors, Handbook of Environmental Economics Vol. II. Amsterdam: North Holland.

Randall, Alan. 1994. A Difficulty with the Travel Cost Method. Land Economics 70(1):88-96.

Rosenberger, R.S. and J.B. Loomis, 2000, "Using Meta-Analysis of Economic Studies: An Investigation of Its Effects in the Recreation Valuation Literature," Journal of Agricultural and Applied Economics 32(3): 459-470.

Smith, V. Kerry and Yoshiaki Kaoru, 1990, "Signals or Noise? Explaining the Variation in Recreational Benefit Estimates," American Journal of Agricultural Economics 72: 419-433.

Walsh, R.G., D.M. Johnson, and J.R. McKean, 1992, "Benefit Trasnfer of Outdoor Recreation Demand Studies, 1968-1988," Water Resources Research 28(3): 707-713

4.3.2 Hedonics

Brief description of the method: Hedonic methods seek to exploit possible relationships between demands for private goods and their associated bundle of characteristics. The logic for the hedonic method stems from a recognition that prices of a set of closely related, but heterogeneous market goods -- types of a car or computer -- reflect the characteristics of each type of the product that describe how it is different. When people select from among the set of available goods, the hedonic model assumes that they will choose the type that is their most preferred given its price and attributes. In equilibrium the set of prices for these differentiated goods will be structured so there is no incentive for anyone to change their choices. The hedonic price function relating prices to characteristics is a reduced form description of this equilibrium condition. The primary applications of this logic in the field of environmental economics involve housing prices and the wage rates for jobs

Assuming that the price of a house reflects the attributes of that house, its property, neighborhood, and facilities that are "near" it, then the hedonic price function reflects a buyer's marginal willingness to pay (WTP) for small changes in one of these attributes. This measure is a single point estimate of the marginal value. The method does not provide the basis for measuring, without additional assumptions, any economic benefits that are associated with a large change in one or more of these attributes. These attributes can include the structural features of the house, its lot, and the characteristics that are conveyed to those living in the home because of its location. For example, if a house is on the coast, residents can experience the coastal views, any beach related amenities, as well as any greater risk of damage that might arise from coastal hazards. If that feature is some aspect of an ecological service available to an individual because she lives in the house, the model allows that incremental value of a change in that service to be estimated.

f the attribute measures a characteristic that can be related to a policy, e.g., proximity to a Superfund site before and after clean-up, then it is possible to describe a buyer's willingness to make tradeoffs for small changes in that attribute. There are important qualifications that must be considered in evaluating the results from these models. For example, to the extent the prices for homes near wetlands or in flood zones are found to be related to (i.e. have a statistically significant association with) the measures that are used to isolate these features, then there is indirect evidence that these features are recognized by buyers and sellers. This result follows because they contribute to the observed equilibrium prices for the homes represented by the hedonic function. Relating such a recognition to a measure of the incremental value for the change in services requires assumptions describing how changes in the variable that can be measured and included in the price function relate to changes in the service of interest.

Extensive data are needed to estimate a statistical function that relates housing prices to housing characteristics that include environmental attributes so that small changes in the quality or quantity of that environmental attribute can be related to small changes in housing prices.

Status as a Method: The hedonic method has been widely used to evaluate site specific amenities and disamenities. Examples of applications involve: air pollution, noise pollution, proximity to water bodies, wetlands, coastal areas, location of homes in hazardous areas such as earthquake or flood zones. See Palmquist (2005) for a general overview of the literature and Smith and Huang (1995) for a meta analysis of the studies of air pollution and property values. This and other meta analyses indicate clear support for the methods for applications where we can expect buyers and sellers to have knowledge of the amenities.

Applications involving site attributes that might be more closely aligned with services of ecosystems are much more limited. Several studies have investigated the effects of proximity to wetlands of different types as well as for distance to open space. Examples include Mahan, et al. (2000), Netusil (2005), and Smith, et al. (2002). An important difficulty in using these results arises in converting the incremental value estimated for a change in distance to a measure more directly related to changes in ecosystem service.

Strengths and Limitations: Hedonic methods are familiar to most people who have purchased or sold a house because realtors do an informal hedonic type analysis comparing homes described as "comparables" to price a proposed new listing.

The main strength of the hedonic housing method is that it is based on people's actual choices. However, all hedonic methods face significant econometric hurdles and are subject to the standard criticism of statistical relationships that they reveal correlation but fall short of revealing causation. Hedonic estimates can be sensitive to the choice of model specification (see, for example, Cropper, Deck and McConnell, 1988). Moreover, relating housing prices to many ecosystem services remains elusive. Finally, hedonic methods assume that consumers operate with full information about local environmental attributes, but this assumption may be unrealistic.

The method is best suited for local housing markets. While several studies have estimated national hedonic property value models, it is generally agreed that it is unreasonable to assume that there is a single national market for housing with an equilibrium that adequately describes the tradeoffs among housing attributes in very different locations. To implement the method for estimating the hedonic price function , it is important to have access to a real estate transaction database with sales prices, housing characteristics, and the latitude/longitude coordinates for each property. These data can then be merged to GIS files describing access to various spatially delineated environmental resources such as air quality as well as to ecosystem services.

Uncertainty: The primary sources of uncertainty with the hedonic model for policy applications arise with the measurement of attributes that are assumed to represent the environmental services available to people due to living in the house. Further research on how people learn about these aspects of a location and what they consider to be conveyed by a location would help to address this issue.

In addition, simulation analysis evaluating the performance of hedonic price functions as approximations to an equilibrium matching process (see would also contribute to our understanding of the sensitivity of the method to assumptions about model structure and functional form. See, for example, Cropper, Deck and McConnell (1988)

Key References:

Champ, P.A., K.J. Boyle and T.C. Brown, editors. 2003. A Primer on Non-Market Valuation. Dordrecht: Kluwer Academic Press.

Cropper, M.L, L. Deck, and K.E. McConnell, 1988, "On the Choice of Functional Forms for Hedonic Price Functions," Review of Economics and Statistics, 70: 668-75.

Freeman, A.M. III. 2003. The Measurement of Environmental and Resource Values, second edition. Washington, D.C.:Resources for the Future).

Haab, T.C. and K.E. McConnell. 2002. Valuing Environmental and Natural Resources, Cheltenham, UK: Edward Elgar.

Mahan, B.L., S. Polasky, and R.M. Adams, 2000, "Valuing Urban Wetlands: A Property Price Approach," Land Economics, 76 (February): 100-113.

Netusil, Noelwah, 2005, "The Effect of Environmental Zoning and Amenities on Property Values: Portland Oregon," Land Economics, 81 (May): 227.

Palmquist, Raymond B., 2005, "Hedonic Models" in K. Mäler and J. Vincent, editors, Handbook of Environmental Economics Vol. II Amsterdam: North Holland.

.V. Kerry Smith, Poulos, Christine, and Hyun Kim, 2002, "Treating Open Space as an Urban Amenity," Resource and Energy Economics 24: 107-129.

4.3.3 Averting behavior models

Brief Description of the Method: Averting or mitigating behavior models simulate consumer behavior and rely on the existence of an activity that substitutes for the services provided by an environmental resource. The averting behavior method infers values from "defensive," mitigating, or "averting" expenditures, i.e. those actions taken to prevent or counteract the adverse effects of environmental degradation. By analyzing the expenditures associated with these defensive purchases, researchers impute a value that individuals place on small changes in environmental or health risks. In effect, a defensive expenditure is spending on a good that is a substitute for health protection or an environmental quality or service. Because the method is based on an estimation of the marginal rate of technical substitution between the environmental service and a market good or service with a known market price, it is capable of producing monetary estimates of the value of the environmental service. What is required is an understanding of the technical relationships underlying the

ability of the environmental service and its market good substitute to enhance human well-being.

Status of the Method: There is a substantial literature on the theoretical dimensions of the method (for example, Freeman 2003, Dickie, 2003, Smith, 1991) but relatively few convincing studies demonstrating it will work in practice. Examples of defensive expenditures include the choice of automobile type (as it relates to fatality risk), safety helmets, fire alarms, and water filters. However, since these expenditures only capture a portion of an individual's willingness to pay (WTP) for these protections, averting behavior results are sometimes interpreted as a lower bound on willingness to pay to avoid a particular harm. The most common application of averting behavior models has been the estimation of values for morbidity (illness) risk. Averting behavior studies rarely provide economic values for ecosystem services. Even for those averting behavior studies for water quality, the motivation for the averting behavior is usually to protect health or life.

Key References:

Dickie, Mark. 2003. "Defensive Behavior and Damage Cost Methods," in Champ, P.A., K.J. Boyle and T.C. Brown, editors, A Primer on Non-Market Valuation. Dordrecht: Kluwer Academic Press.

Freeman, A.M. III 2003 The Measurement of Environmental and Resource Values, second edition .Washington, D.C.: Resources for the Future.

Smith, V. K. (1991), "Household Production Functions and Environmental Benefit Estimation," in J. B. Braden and C. D. Kolstad, eds., Measuring the Demand for Environmental Quality. Amsterdam: North Holland

4.4. Non-market Methods – Stated Preference

Brief Description of the Method: Stated preference methods rely on surveys which ask individuals to make a choice, describe a behavior, or to state directly what they would be willing to pay for specified changes in environmental services not traded in markets. The various stated preference techniques are distinguished by how the information is presented, what questions are asked, and how their responses are formatted. It is important to acknowledge that the choices, stated values, or revised patterns of use are derived from answers to questions that ask respondents what they would do, or how much they would pay

for, or how they would alter their choices in response to changes in the amount of a non-market good or service in a specified hypothetical setting. This is in contrast to Revealed Preference Methods which are based on observing the actual choices made by people facing real constraints on income etc.. Stated preference methods offer the opportunity to measure tradeoffs for anything that can be presented as a credible and consequential choice and hence their primary advantage is flexibility.

Although not all authors use the same terminology, the term stated preference methods has come to refer to any survey based study in which respondents are asked hypothetical questions that are designed to reveal information about their preferences or values. The term encompasses three broad types of questions. The first type involves questions that ask directly about monetary values for a specified commodity or environmental change. These are usually called contingent valuation questions (CVM). In the past the most commonly used CVM questions simply asked peo¬ple what value they place on a specified change in an environmental amenity or the maximum amount they would be willing to pay to have it occur. These are usually open-ended in that the individual has to state a number rather than respond to a number offered by the researcher. The responses to these questions, if truthful, are direct expressions of value. The other major type of CVM question asks for a yes or no answer to the question, “Would you be willing to pay $X ...?" Each individual's response reveals only an upper bound (for a no) or a lower bound (for a yes) on the relevant welfare measure. Questions of this sort are termed discrete choice questions. Responses to discrete choice questions can be used to estimate willingness to pay functions or indirect utility functions.

The second and third major types of Stated Preference methods do not reveal monetary measures directly. Rather, they require some form of analytical model to derive welfare measures from responses to questions. The second type of question is called variously "choice experiment," "conjoint analysis," or sometimes a "choice experiment" or an "attributes based method (Holmes and Adamowicz, 2003). In this approach to questioning respondents are given a set of hypothetical alterna¬tives, each depicting a different bundle of environmental attributes. Respondents are asked to choose the most preferred alternative, to rank the alternatives in order of preference, or to rate them on some scale. Responses to these questions can then be analyzed to determine, in effect, the marginal rates of substitution

between any pair of attributes that differentiate the alternatives. If one of the other characteristics has a mone¬tary price, then it is possible to compute the respondent's willingness to pay for the attribute on the basis of the responses. Studies based on this form of question

In the third type of SP question, individuals are asked how they would change the level of some activity in response to a change in an envi¬ronmental amenity. If the activity can be interpreted in the context of some behavioral model such as an averting behavior model or a recreation travel cost demand model, the appropriate indirect valua¬tion method can be used to obtain a measure of willingness to pay. These are known as contin¬gent behavior or sometimes contingent activity questions.

Since most of the questions about the status of the method and strengths and limitations are common to the three question formats, they will be discussed together in the remainder of this section.

Status of the Method: The method has an extensive literature of principles and applications extending over a forty year period. Mitchell and Carson's (1989) pioneering treatise is still the primary reference on CVM, especially for design and implementation questions. See also Carson (1991). Two new works that focus on best practice and empirical estimation for CVM and stated choice studies are Boyle (2003) and Holmes and Adamowicz (2003), respectively. The so-called NOAA Blue Ribbon Panel (U.S. National Oceanic and Atmospheric Administration 1993) reviewed CVM in the context of assessing damages to natural resources in support of litigation and provided its guidelines for best practice. Other important references are: Bjornstad and Kahn (1996) for a review of theoretical and empirical issues that includes assessments by both proponents and critics of stated preference methods; Kopp, et al. (1997); Bateman and Willis (1999); Bateman, et al., (2002) and Smith(2004,2007).

Use of the stated preference methods for environmental valuation has been controversial. This controversy had the salutary effect of stimulating a substantial body of new research on both practice and on the credibility or validity of stated preference estimates of value. A good overview of the issues raised in this controversy is contained in the three essays published as a symposium in the Journal of Economic Perspectives (Portney 1994, Hanemann 1994, and Diamond and Hausman 1994). See also, Hausman (1993).

A major issue concerning the status of stated preference methods is the validity of the resulting value estimates. There are several concepts of validity and various approaches to assessing the validity of responses. See Freeman (2003) and references therein for further discussion.

This probably would be a good place to insert something about Paul Slovic's view of "preference construction.- Rick F.

Another issue is the degree of incentive compatibility of the various questioning formats and the set of methods as a whole. Carson, et al. (2000), reasoning from first principles about what is in the best interest of respondents faced with a scenario, payment vehicle, and elicitation question, have established under what conditions stated preference questions give people incentives to reveal their true values.. The first two conditions are that the survey question be about something that matters to the respondent and the respondent believe that his/her response might affect the outcome of the policy issue that is the subject of the survey. If both conditions hold, then the survey question is termed “consequential” to respondents. For consequential questions, it is possible to reason from an assumption of acting on rational self interest to predict whether responses will be truthful and if not, then at least in some cases what the direction of bias will be.

For consequential questions, the only question format that can in principle be incentive compatible is the single discrete choice question. And this form requires the further condition that the government agency is perceived as being able to compel payment of some amount from the respondent if the good is provided. For example, questions that ask about the willingness to make a voluntary contribution to support some government action fail this condition and provide incentives to respond “yes” even when the requested contribution is greater than the respondent’s WTP.

<u>Strengths and Limitations</u>: Strengths include the accumulated experience of forty years of practice and research. Also in principle, stated preference methods are the only set of methods capable of capturing so-called nonuse values, since without use there is no behavior that can reveal values through application of revealed preference methods.

A limitation is that since responses to questions must reflect in some sense the knowledge that individuals have about the thing being valued as well as respondents'

preferences, the methods cannot be used to value ecosystem services about which people are ignorant. For example if respondents were asked questions concerning phytoplankton but were ignorant of the role or phytoplankton in supporting the aquatic food chain and higher order species that they might value, their responses might be interpreted as placing no value on phytoplankton. One solution to this problem is to use the survey instrument to convey information to respondents about the role of the ecosystem service being valued and the potential consequences of changes in the level of this service. See for example Banzhaf, et al. (2004). Then, of course, the question becomes one of the validity of the information provided to respondents and the potential for biasing responses by providing biased information.

Key References

Banzhaf, Spencer, Dallas Burtraw, David Evans, and Alan Krupnick. 2004. Valuation of Natural Resource Impprovements in the Adirondacks. Washington, DC: Resources for the Future.

Bateman, Ian J., et al. 2002. Economic Valuation with Stated Preferences: A Manual. Cheltanham, UK: Edward Elgar.

Bateman, Ian J., and Kenneth G. Willis (eds.). 1999. Valuing Environmental Preferences: Theory and Practice of the Contingent Valuation Method in the US, EU, and Developing Countries. Oxford, UK: Oxford University Press.

Bjornstad, David J., and James R. Kahn (eds.). 1996. The Contingent Valuation of Environmental Resources: Methodological Issues and Research Needs Cheltenham, UK: Edward Edgar.

Boyle, Kevin J. 2003. Contingent Valuation in Practice. In A Primer on Non-market Valuation edited by Kevin J. Boyle and Patricia A. Champ. Boston: Kluwer Academic Publishers.

Carson, Richard T. 1991. Constructed Markets. In Measuring the Demand for Environmental Quality, edited by John Braden and Charles Kolstad. Amsterdam, The Netherlands: Elsevier.

Carson, Richard T., Theodore Groves, and Mark J. Machina. 2000. Incentive and Informational Properties of Preference Questions. Unpublished. http;//weber.ucsd.edu/~rcarson/ (accessed on August 20, 2002).

Diamond, Peter A., and Jerry A. Hausman. 1994. Contingent Valuation: Is Some Number Better Than No Number? Journal of Economic Perspectives 8(4):45-64

Freeman, A. M. III. 2003. The Measurement of Environmental and Resource Values: Theory and Methods. 2nd ed. Washington, DC: Resources for the Future

Hanemann, W. Michael. 1996. Valuing the Environment Through Contingent Valuation. Journal of Economic Perspectives 8(4):191-43.

Hausman, Jerry A., (ed.). 1993. Contingent Valuation: A Critical Assessment. Amsterdam: North-Holland.

Holmes, Thomas, and Wictor Adamowicz. 2003. Attribute-Based Methods. In A Primer on Non-market Valuation edited by Kevin J. Boyle and Patricia A. Champ. Boston: Kluwer Academic Publishers.

Kanninen, Barbara J. editor 2007 Valuing Environmental Amenities Using Stated Choice Studies (Dordrecht,The Netherlands: Springer)

Kopp, Raymond J., Werner W. Pommerehne, and Norbert Schwarz (eds.). 1997. Determining the Value of Non-Marketed Goods: Economic, Psychological, and Policy Relevant Aspects of Contingent Valuation Methods. Boston: Kluwer Academic Publishers.

Mitchell, Robert Cameron, and Richard T. Carson. 1989. Using Surveys to Value Public Goods: The Contingent Valuation Method. Washington, D.C.: Resources for the Future.

Portney, Paul R. 1994. The Contingent Valuation Debate: Why Economists Should Care. Journal of Economic Perspectives 8(4):3-17.

Smith, V. Kerry 2004, “Fifty years of Contingent Valuation” in T. Tietenberg and H. Folmer editors,International Yearbook of Environmental and Resource Economics 2004/2005 (Chetlenham,U.K. Edward Elgar) pp1-60.

Smith, V. Kerry 2007 “Judging Quality” in B. Kanninen editor Valuing Environmental Amenities Using Stated Choice Studies (Dordrecht,The Netherlands: Springer)

U. S. National Oceanographic and Atmospheric Administration. 1993. Report of the NOAA Panel on Contingent Valuation. http://web.lexis-nexis.com/congcomp/printdoc (accessed on August 12, 2002).

4.4.1 Combining Revealed and Stated Preference Methods

It is possible to combine revealed and stated preference methods to estimate what both types of choices imply for characterizing an individual's willingness to pay for changes in environmental services. Cameron (1992) was the first to propose this idea for environmental applications. To be informative this strategy must be based on an analysis of the revealed and stated behaviors to establish that the empirical models used to describe these outcomes share at least one parameter. That is they must each be capable of identifying at least one common parameter. Ideally there would be more parameters shared between the models. Most applications collect the two types of data (i.e. revealed and stated preference) from the same respondents. This requirement is not essential. It would be possible in principle to combine samples with different respondents providing the revealed and stated components of the analysis. A key issue in applying these methods to the task of valuing ecosystem services is the need to have measures for the quality and amount of ecosystem services that are compatible with models and data typically available for revealed and stated preference models.

See Adamowicz, et al., (1994), Earnhart (2001, 2002), and McConnell, et al. (1999) for more recent applications.

Key References:

Adamowicz, W., J. Louviere, and M. Williams. 1994. Combining Revealed and Stated Preference Methods for Valuing Environmental Amenities. Journal of Environmental Economics and Management 26(3):271-292.

Cameron, Trudy A. 1992. Combining Contingent Valuation and Travel Cost Data for the Valuation of Nonmarket Goods. Land Economics 68(3):302-317.

Earnhart, Dietrich. 2001. Combining Revealed and Stated Preference Methods to Value Environmental Amenities at Residential Locations. Land Economics 77(1):12-29.

Earnhart, Dietrich. 2002. Combining Revealed and Stated Data to Examine Housing Decisions Using Discrete Choice Analysis. Journal of Urban Economics 51(1):143-169.

McConnell, Kenneth E., Quinn Weninger, and Ivar E. Strand. 1999. Joint Estimation of Contingent Valuation and Truncated Recreational Demands. In Valuing Recreation and the Environment: Revealed Preference Methods in Theory and Practice, edited by Joseph A. Herriges and Catherine L. Kling. Cheltanham, UK: Edward Elgar.

4.5. Ecosystem Benefit Indicators

Brief Description: Ecosystem benefit indicators (EBIs) offer a quantitative but not monetary approach to the weighting of ecological characteristics and changes. They can also serve as inputs to benefit function transfer analysis, which is a monetary weighting technique. EBIs provide a way to illustrate ecological benefits in a specific setting. For example, if water is available at a particular place and time, how many water users (e.g., recreators, farms) are present to enjoy that service? What other sources of water are available to those same user? These questions are central to economic valuation of the resource. EBIs are countable landscape features that tell us about demand for, scarcity of, and complements to particular ecosystem services.

Key inputs: EBIs are drawn mainly from geospatial data, including satellite imagery. Data can come from state, county, and regional growth, land-use, or transportation plans; federal and state environmental agencies; private conservancies and nonprofits; and the U.S. Census.

Key outputs: Spatially specific measures (both geo-coded data and corresponding visual depictions) of social and biophysical features that enhance or decrease the desirability of particular ecosystem services.

Scale: The method is entirely scalable. A strength, however, is the ability to relate ecological and economic features in a specific landscape context. For example, the method can be applied to individual projects, investments, or decisions made in a particular watershed. They can also be expressed as local, regional, state, or national aggregates.

Example of how the method could be used as part of the CVPESS framework:

The method relates to framework item (4): "Characterization of the Value of Changes in Monetary and Non-Monetary Terms."

Benefit indicators are countable features of the physical and social landscape. More specifically, they are features that influence – positively or negatively – ecosystem services' contributions to human wellbeing. The consumption of services often occurs over a wide scale. For example, habitat support for recreational and commercial species, water purification, flood damage reduction, crop pollination, and aesthetic enjoyment are all services typically enjoyed in a larger area surrounding the ecosystem in question. EBIs help us understand the larger social and physical landscape so that we can better assess the relative importance of particular services in particular places at particular times.

Economic measures of benefits are affected by the following: the ecosystem feature's scarcity, natural and built substitutes, complementary inputs, and the number of people in proximity to it. For a given ecosystem service scarcity, substitutes, complements, and demand can be related to landscape characteristics. Landscape features that relate to human wellbeing can be systematically counted and mapped, then aggregated into bundles of indicators (an index). Some indicators are biophysical, others relate to the socio-economic environment.

Benefit indicators are an input to a wide variety of tradeoff analysis approaches, but do not themselves make or calculate the results of such tradeoffs. First, they can be used as ends in themselves as regulatory or planning performance measures. Second, they can be used as part of public processes designed to communicate the implications of a change or policy across a variety of scales. Indicators or an index based on them can then be used to elicit public preferences over environmental and economic options – as in mediated modeling exercises or more informal political derivations. In this way, benefit indicators are a potentially powerful complement to group decision processes. Third, they can be used as inputs to economic and econometric methods such as benefit transfer, or stated preference models. This is an area where research is needed. Economic methods must be developed to link indicator outcomes to dollar-based valuation in a way that is both statistically and theoretically sound. In principle, benefit indicators could be used to calibrate the transfer function in benefit transfers. They could also be be used to systematize alternative choice scenarios in choice experiments and stated preference surveys.

As a method to inform the weighting of ecosystem services in a social decision context, the benefit indicators method requires information provided by the biophysical sciences. The method requires spatially depicted biophysical endpoints. EBIs are then related to those endpoints.

The method can be applied to any ecosystem service benefit where benefits are related to the spatial delivery of services and social landscape in which the benefit is enjoyed. Existence benefits (where spatial location is irrelevant to both provision and value) are the only ecosystem benefit category where the method would be inapplicable.

The data used in EBI analysis is well-suited to delivery via a national data bank.

Status as a Method: The method is new and thus relatively undeveloped. EPA has funded a small amount of research on the topic. For citations to peer reviewed research, see below.

Strengths/Limitations: EBIs are designed to be a relatively non-technical way to express the factors that contribute to conventional economic measures of benefits provided by ecosystem services. Their simplicity, and transparency, is an advantage.. They can be used to communicate and educate. By stopping short of monetary estimation of benefits (unless integrated in a benefit function transfer method) they are also a way for the agency to overcome resistance to economic assessments of the natural world – while still conveying outcomes in a way designed to be consistent with economic principles and the dependence of human well-being on natural assets.

The principle disadvantage is that they do not directly yield dollar-based ecological benefit estimates. They also do not in themselves weight or estimate the tradeoffs associated with different factors relating to benefits (though as noted above they can be married to more formal methods designed to do such weighting).

Because indicators can be cheaper to generate than econometric value estimates they better allow for landscape assessment of multiple services at large scales.

Treatment of Uncertainty: A core rationale for the use of a benefit indicator approach is to explicitly convey the sources of complexity – and hence uncertainty – characterizing biophysical systems and the benefits arising from them. The visual depiction of benefit indicators, for example, can mimic sensitivity analysis by presenting a range of benefit scenarios in GIS form. However, the visual depiction of quantitative information introduces

uncertainties of its own. In particular, visual depictions can strongly influence perceptions. Uncertainty with regard to how indicators are perceived, particularly when presented visually should be acknowledged.

Research Needs

- Integration of EBIs with biophysical endpoints
- Integration of EBIs with econometric valuation methods (benefit function transfer, stated preference and choice modeling)
- Suitability for group decision techniques, such as mediated modeling
- Practical application to illustrate data needs and measurement issues
- Satisfying these needs would be a significant undertaking in terms of expertise, financial resources, and coordination within the agency.

References

James Boyd, "What's Nature Worth? Using Indicators to Open the Black Box of Ecological Valuation," Resources, 2004.

James Boyd and Lisa Wainger, "Landscape Indicators of Ecosystem Service Benefits," 84 American Journal of Agricultural Economics, 2002.

James Boyd, Dennis King, and Lisa Wainger "Compensation for Lost Ecosystem Services: The Need for Benefit-Based Transfer Ratios and Restoration Criteria," 20 Stanford Environmental Law Journal, 2001.

Lisa Wainger, Dennis King, James Salzman, and James Boyd, "Wetland Value Indicators for Scoring Mitigation Trades," 20 Stanford Environmental Law Journal, 2001.

4.6. Transfer of Economic Information

Text to be inserted here

5. DELIBERATIVE PROCESSES FOR ELICITING VALUES

Introductory text needed.

5.1. Valuation by Decision Aiding/Structured Decision Making

The committee uses the term "value" in a broad sense that includes moral, spiritual, religious, aesthetic, and other attributes appropriate considerations to guide choices and policies. Values are normative, which means that they shape thinking and preferences about what one one ought to do and how one ought to feel. In many contexts values themselves are shaped by processes of deliberating about different actions, policies, and states of affairs and comparing them to each other.

Deliberative methods for valuation need to be developed in a way that is sensitive to the insights provided by more than 30 years of research on decision making. One of the main themes that emerged from this research is that people's preferences and values for complex, unfamiliar goods, are often constructed during the process of elicitation (Lichtenstein & Slovic, 2006). In keeping with the constructive nature of environmental preferences and values, designers of a valuation exercise must function not only as archeologists, carefully uncovering what values exist, but also as architects, working to build a defensible expression of value (Gregory, Lichtenstein, & Slovic, 1993).

Text Box 10: The Constructive Nature of Environmental Preferences

One achievement of behavioral research on decision making has been to demonstrate the strong influence of context on measures of preference and value. As a result, decision scientists seeking to elicit values have recognized that order of effects, the range and mixture of items being evaluated, the amount and nature of information provided about each item, the method of elicitation, and many other contextual factors can affect the results of any serious elicitation attempt. The nature of these effects are not merely "biases" to be minimized by more careful procedures. Rather, they represent fundamental changes in the preferences and values themselves.

Research on what has come to be known as "the constructive nature of preference" has strong implications for attempts to value environmental impacts. Consider, for example, the phenomenon of preference reversal, which has been studied by psychologists and

economists for more than twenty-five years. For example, Irwin et at. (1993) conducted several studies showing preference reversals in willingness to pay responses (WTP). These studies involved comparisons of improvements in consumer goods, such as a better camera, with improvements in air quality. They found that WTP based on a single-stimulus response favored improvements in consumer goods and WTP based on a choice response (e.g. which improvement would you pay more to obtain?) favored improvements in air quality. Findings of preference reversals involving environmental values provide strong evidence for the constructive nature of preference. As Tversky, Sattath, and Slovic (1988) observed:

> In the classical analysis, the relation of preference is inferred from observed responses and is assumed to reflect the decision maker's underlying utility or value. But if different elicitation procedures produce different orderings of options, how can preferences and values be defined? And in what sense do they exist? (p. 383)

An important corollary of the constructive view is that the strong values that people hold for environmental goods are not represented in their minds in monetary form. Consider all the goods that we might want to value in dollar terms. These could be arrayed on a continuum according to the level of market experience that we have with them. At one extreme, would be' goods such as groceries, for which market experience is great and the strength of our values or preferences can be relatively easily represented by a market price. As we move from groceries to appliances, automobiles, and homes, market experience lessens and the ease of representing our preferences monetarily declines as well. For goods such as air or water quality, wilderness areas, endangered species, and many other elements of the natural environment, the market no longer applies, and the link between values and money becomes tenuous-so tenuous that it may not exist. Thus, we can have strongly held values that are not at all coded mentally in terms of dollars. Attempts to translate such values into monetary equivalents must take special cognizance of this problem.

<u>Brief description of the method.</u> Individual valuation methods must be responsive to a set of identified objectives. For example, if an objective in a given decision is to improve environmental health, desired attributes that will define “environmental health” must first be

identified and operationalized. Decision makers, analysts, and stakeholders must first establish what "environmental health" will mean as well as the attributes of it that will guide its measurement. In a decision making context, the first aspect of measuring value, therefore, entails predicting the level of achievement that can be realized by implementing one of a set of management options across the objectives and their associated system-specific attributes that come to define their objectives of interest (simply "environmental health" in the example above). In an evaluative context, this first step entails measuring the level of change—improvement or degradation—that has be realized, relative to some previous state, across the objectives and their associated attributes that are of interest. These processes of measurement or prediction involve the application of the various analytic methods within the specific domain of interest.

Once objectives and their system-specific attributes have been specified and values obtained across the full range of alternatives under consideration, a decision maker (or decision makers depending upon whether the choice calls for judgments by an individual or a group) can go about the process of:

- selecting or identifying a suitable management option (in a decision making context), or
- making judgments about the current state of the system relative to a previous state (in the case of evaluation).

This process, by necessity, involves confronting tradeoffs when attributes of value or objectives inevitably conflict. The overall value of the benefits obtained from ecological systems and services are ultimately reflected in: a) the full spectrum of values obtained across all of the stated objectives and their associated attributes and b) the tradeoffs that are made when:

- in the case of decision making, comparing one alternative—which yields a suite of benefits and costs—with others, or
- in the case of evaluation, comparing the improvements or declines that are presently realized relative to some past state.

Text Box 11: Structuring the Valuation Problem

In the important first step of the valuation process, the analyst collects, lists, and organizes a description of the problem, identifying all the attributes (that is, all the aspects of the problem that have value to people). The goal is to develop an explicit, comprehensive picture of all factors that contribute significantly to the value of the good or activity. To do so, the analyst will consult both technical experts, to get the facts, and the affected citizenry, to find the link between the facts and the values.

Diverse groups of people should be consulted to select the value attributes. These stakeholders are defined in an operational sense as groups of people who, for any reason (e.g., place of residence, occupation, favored activities), share common concerns or opinions regarding a proposed action. The analyst might convene several stakeholder groups, each composed of seven to ten; from each group, a values structure is elicited. Careful selection of stakeholder groups ensures that the full range of views is adequately covered. For example, the representatives of an environmental advocacy organization might be expected to present a somewhat different list of attributes than would members of the local Chamber of Commerce, but the views of each group are likely to encompass those of many other citizens.

Suppose that someone wanted to construct a value for the damage resulting from a specific pollutant accidentally spilled into an estuary. Technical experts can provide information describing the lake before and after the damage. These descriptions then can be presented to representatives of the people affected by the damage to identify the value attributes. For example, the physical event of the death of a large number of fish might imply aesthetic loss (when the dead fish wash up on the shore), loss of genetic diversity, and loss of commercial fishing jobs and profits. These losses indicate the value attributes.

Generic attributes for the lake problem might be Effects on Scenic Beauty, Effects on Genetic Diversity, Human Health Effects, Effects on Commerce, and so forth. Each attribute would have sub-attributes. For example, sub-attributes influencing Effects on Commerce might be Real Estate Values (the price of vacation homes would go down if the shore line becomes ugly), Tourist Values, and Entitlement (expressing the general public's non-use value for a beautiful lake). Some or all of these sub-attributes might be further broken down into sub-sub-attributes, and so forth, until all relevant values have been listed and organized. At the lowest level, each attribute is described in terms of some specific measure. For

example, one sub-component of Scenic Beauty concerned with shoreline attractiveness might have as its bottom-level measure the number of dead fish per acre of beach.

After the structure of the problem has been determined, deliberation and argument can then be organized around the identified multidimensional attributes of value. In some circumstances, if monetization is required, the techniques of multiattribute utility theory can be used to construct a monetary value based on the identified attributes of value (Box 3; see, Gregory, Lichtenstein, & Slovic, 1993).

Text Box 12: How to Monetize if You Must

If the overall ecosystem benefit of some action must be expressed in monetary form, due to legal requirements, the multiattribute method can be extended to do this. The first step is to elicit utilities (values) from each group of respondents.

Utilities are assessed for every lowest-level value scale. To start with, it is convenient to assess every utility on a common scale, say, from 0 to 100. For example, the maximum number of dead fish per unit area as a result of a pollutant spill might be assigned a utility score of 100, and the minimum impact level, perhaps 0 dead fish per acre, would be assigned a score of 0. It is essential that this range of outcomes be carefully specified, that the range encompasses all reasonably possible values of the attribute measure, and that this range, once chosen, remain fixed throughout the analysis. Trade-offs are then assessed, using weights and multiplicative factors that rescale the various utilities in recognition that not all attributes of value are equally important.

All these value elicitations would be done with numerous consistency checks. If you have told the analyst that a change from 0 to 100 on scale A is twice as good as a change from 0 to 100 on scale B, and that a change from 0 to 100 on scale B is four times as good as such a change on scale C, the analyst will then check to see that you do, indeed, believe that the scale A change is eight times as good as the scale C change.

Once all the pieces are in place, a combination rule specifies how to calculate the total utility for any particular plan, program, or scenario. This total utility will be expressed using a single arbitrary “utile” unit of measurement. Thus unit would be dollars. In theory, this conversion need only be made at one place in the model. For example, one such conversion

might be made in the estuarine pollution example by noting the monetary value and the utility (or, here, disutility) of the loss of one ton of fish to one fishery. Because all utilities, including this one, are measured on a common scale, the monetary worth of all utilities, including the total utility, can be computed from this one conversion. In practice, of course, one would want to find several parts of the model for which both utilities and their monetary equivalents are known (e.g., real estate values, perhaps even the value of a life.)

Highlights of method relative to major C-VPESS recommendations:

Is value monetized or not? Not necessarily; it depends upon the decision context and the agreed upon objectives, along with the attributes and measures associated with these objectives, that determines how the value of (a) alternatives or (b) present as compared to past states will be established.

Does the method help to identify ecosystem service impacts of EPA action or uses ecosystem service information? Yes.

Does the method identify changes in ecosystems and their services that are likely to be of greatest concern to people? Potentially; it depends upon who is involved in the process of defining the decision problem and eliciting objectives (and the associated attributes and measures) for it.

Does the method address ecological impacts that wouldn't normally be included by EPA or are not easily addressed by monetized approaches? Yes. This method can incorporate the full spectrum of attributes that contribute to value.

If the method does not present values that are in monetizable economic units; does it present some basis or rationale for judging or characterizing the importance of alternative expressions of value? Yes; these "alternative expressions" are reflected in a systematic evaluations of: a) the full spectrum of values obtained across all of the stated objectives and their associated attributes and b) the tradeoffs that are made when:

- in the case of decision making, comparing one alternative—which yields a suite of benefits and costs—with others, or
- in the case of evaluation, comparing the improvements or declines that are presently realized relative to some past state.

Does the method involve interdisciplinary collaboration among physical/biological and social scientists; and ecologists at an early stage? Potentially; it depends upon context of the decision problem and the objectives that must be met when making a decision. There is nothing in the method that precludes such collaboration.

Does the method involve the public in identifying values at an early stage and allow the public recurring involvement through the process? Potentially. This is quite likely to be desirable and the method allows public input into the construction of value.

Are there types of data, and information that could be generated through use of the method that might be useful to have in a data/model bank so that valuation can be done more easily by EPA in the future? It is unlikely that this approach would benefit from the creation of a data/model bank; however, past applications of this approach—especially when it is implemented for similar kinds of problems in a particular region—have benefited from objectives and attributes (vs. benefits) transfer. In other words, many of the objectives that are important to stakeholders and the system-specific attributes associated with them may be the same across different decision/evaluation problems.

Status as a method. This method, applied under the banner of decision aiding or decision structuring, has been extensively applied both experimentally and through practical case studies. Results from many of these efforts, as well as their theoretical foundations, have been peer reviewed and/or published (Arvai & Gregory 2003; Arvai et al. 2001; Gregory 2000; Gregory et al. 2001; Gregory & Failing 2002; Gregory et al. 1993; Gregory, Lichtenstein, & Slovic 1993; Keeney & McDaniels 1992; Keeney 1992; Keeney & Gregory 2005; Keeney & Raiffa 1993; Lichtenstein & Slovic 2006; McDaniels et al. 1999; McDaniels & Trousdale 1999; Payne et al. 1993; Payne et al. 1992; Payne et al. 1999; Slovic 1995). The committee is unaware of any EPA applications.

This approach has been applied in a wide variety of contexts; e.g., informing the development of a national energy policy, relicensing hydroelectric facilities, and environmental restoration. These cases use methodologies developed by decision analysts and behavioral decision theorists.

Strengths/Limitations. The tradeoff judgments required as inputs to this method will not be easy ones to make. But they are not holistic judgments, requiring the simultaneous integration of many dimensions of value. Thus, it is less likely that important aspects of

value will be lost because of cognitive overload. Most importantly, the identified values ideally will be elicited from a wide range of the potentially affected stakeholders. These stakeholders have a right to express their objectives and concerns, thereby contributing to the valuation process, as part of an open and constructive decision-aiding process.

Past experience suggests that the results from these methods are both more easily understood and met with higher levels of stakeholder support when compared with "traditional" (i.e., single decision maker, unimetric) approaches.

The chief advantage of this method is that it promotes and results in defensible, credible evaluations and decisions. The method has the potential to work well under all decision and evaluation contexts. It is transparent, allowing interested parties to see exactly what attributes contribute to value and how they are being assessed and weighted.

Every value-elicitation method affects the values being elicited. We cannot know the exact effects this method would exert on people's values. But the process and results of this method would be explicitly recorded and thus open to scrutiny and critique.

The method can be applied in many EPA decision-making contexts (e.g., rulemaking, site-specific decisions, collaborative decision-making, retrospective evaluation). It would work poorly when the objectives and their system-specific attributes are weakly or incompletely defined because it would result in tradeoffs and hence, valuations that proceed without a thorough understanding of the elements that contribute to value.

<u>What scientific, moral, legal, or psychological obstacles are there to the effective use of the method?</u> If the scientific understanding of impacts is lacking, the method cannot be effective. The method could potentially overcome the objection to other methods that fail to include moral and other non-monetary aspects of value. The method gives standing to these aspects. There are no legal obstacles to use of the method. The method needs to be implemented under the supervision of a skilled analyst, one who understands how contextual nuances can influence the value construction process. Such understanding should guide attempts to minimize biases in valuation.

<u>How resource-intensive is the application of the method?</u> The most important resources are: a) time, b) expertise in the decision sciences, and c) commitment on the part of the agency or group convening the process. Of course, all of these amount to dollars. However, just as EPA has on staff a group of economists, there's no reason why some of

those FTEs could not be converted into decision science positions. Also of importance, as analysts and participants become more familiar with the approach, the time required to complete a given analysis would likely be reduced.

What practical obstacles (e.g., institutional, data) are there to the effective use of this method? There may be significant institutional barriers to the full adoption at EPA of this method (the method is "overly" transparent and frequently highlights objectives and/or alternatives that may not be favored by managers).

What other significant questions may be associated with this method and how would these be addressed? As with most other methods, the participants in the value construction exercise need to be selected. They need to be knowledgeable and representative of the interested and affected parties.

Treatment of uncertainty. This method does incorporate a unique approach for treating uncertainty. Tradeoffs (when evaluating options or undertaking evaluations) can be made across a probability distribution of predicted benefits or outcomes. Also, this method has, in the past, been combined with so-called "open-loop" approaches that proceed based on the need to reopen, evaluate, and sometimes re-make decisions over time (as more information about a managed system is obtained).

Research needs. This method does incorporate a unique approach for treating uncertainty. Tradeoffs (when evaluating options or undertaking evaluations) can be made across a probability distribution of predicted benefits or outcomes. Also, this method has, in the past, been combined with so-called "open-loop" approaches that proceed based on the need to reopen, evaluate, and sometimes re-make decisions over time (as more information about a managed system is obtained).

Key References

Arvai, J., and R. Gregory. 2003. A decision focused approach for identifying cleanup priorities at contaminated sites. Environmental Science & Technology **37**:1469-1476.

Arvai, J. L., R. Gregory, and T. McDaniels. 2001. Testing a structured decision approach: Value-focused thinking for deliberative risk communication. Risk Analysis **21**:1065-1076.

Gregory, R. 2000. Using stakeholder values to make smarter environmental decisions. Environment **42**:34-44.

Gregory, R., J. L. Arvai, and T. McDaniels. 2001. Value-focused thinking for environmental risk consultations. Research in Social Problems and Public Policy **9**:249-275.

Gregory, R., and L. Failing. 2002. Using decision analysis to encourage sound deliberation: Water use planning in British Columbia, Canada. Journal of Policy Analysis and Management **21**:492-499.

Gregory, R., S. Lichtenstein, and P. Slovic. 1993. Valuing environmental resources: A constructive approach. Journal of Risk and Uncertainty **7**:177-197.

Keeney, R., and T. McDaniels. 1992. Value-focused thinking about strategic decisions at B.C. Hydro. Interfaces **22**:94-109.

Keeney, R. L. 1992. Value-focused Thinking. A Path to Creative Decision Making. Harvard University Press, Cambridge, MA.

Keeney, R. L., and R. Gregory. 2005. Selecting attributes to measure the achievement of objectives. Operations Research **53**:1-11.

Keeney, R. L., and H. Raiffa 1993. Decisions with multiple objectives: Preferences and value tradeoffs. Cambridge University Press, Cambridge, UK.

Lichtenstein, S., and P. Slovic. (Eds.). (2006). The Construction of Preference. New York: Cambridge University Press.

McDaniels, T., R. Gregory, and D. Fields. 1999. Democratizing risk management: Successful public involvement in local water management decisions. Risk Analysis **19**:497-510.

McDaniels, T., and W. Trousdale. 1999. Value-focused thinking in a difficult context: Planning tourism for Guimaras, Phillipines. Interfaces **29**:58-70.

Payne, J. W., J. Bettman, R., and E. J. Johnson 1993. The Adaptive Decision Maker. Cambridge University Press, Cambridge, MA.

Payne, J. W., J. R. Bettman, and E. J. Johnson. 1992. Behavioral decision research: A constructive processing perspective. Annual Review of Psychology **43**:87-132.

Payne, J. W., J. R. Bettman, and D. A. Schkade. 1999. Measuring constructed preferences: Towards a building code. Journal of Risk and Uncertainty **19**:243-270.

Slovic, P. 1995. The construction of preference. American Psychologist **50**:364-371.

Tversky, A., Sattath, S., & Slovic, P. (1988). Contingent weighting in judgment and choice. Psychological Review **95**: 371-384.

5.2. Mediated Modeling

Brief description of the method. Computer models of complex systems are frequently used to support decisions concerning environmental problems. To effectively use these models, (i.e. to foster consensus about the appropriateness of their assumptions and results and thus to promote a high degree of compliance with the policies derived from the models) it is not enough for groups of academic "experts" to build and run the models. What is required is a different role for modeling - as a tool in building a broad consensus not only across academic disciplines, but also between science and policy.

Mediated modeling is process of involving stakeholders (parties interested in or affected by the decisions the model addresses) as active participants in all stages of the modeling, from initial problem scoping to model development, implementation and use (Costanza and Ruth 1998; van den Belt 2004). Integrated modeling of large systems, from individual companies to industries to entire economies or from watersheds to continental scale systems and ultimately to the global scale, requires input from a very broad range of people. We need to see the modeling process as one that involves not only the technical aspects, but also the sociological aspects involved with using the process to help build consensus about the way the system works and which management options are most effective. This consensus needs to extend both across the gulf separating the relevant academic disciplines and across the even broader gulf separating the science and policy communities, and the public. Appropriately designed and appropriately used mediated modeling exercises can help to bridge these gulfs. The process of mediated modeling can help to build mutual understanding, solicit input from a broad range of stakeholder groups, and maintain a substantive dialogue between members of these groups. Mediated modeling and consensus building are also essential components in the process of adaptive management (Gunderson, Holling, and Light 1995, van den Belt, 2004).

Example of how the method could be used as part of the CVPESS expanded & integrated framework. As described above, the method is fairly general and could be used to assess any value (means toward an end) that a group of stakeholders could identify and build into a model. Any decision context that requires the estimation of the values of ecosystem goods or services could employ this method, although to the committee's knowledge no EPA

decisions have as yet employed this technique. The method covers all elements of the diagram representing the C-VPESS framework for valuation after the initial identification of EPA needs, and could be used in conjunction with the full range of decision models. Prior applications have been at a broad range of scales, from watersheds or specific ecosystems to large regions and the global scale. The method is in principle broadly applicable to the full range of time and space scales.

- The method is inherently dynamic – that is what it does best
- The results can be aggregated to get a single benefits number as needed.
- Participants in the mediated modeling process gain deep understanding of the process and products. Those who have not participated can easily view and understand the results if they invest the effort. Usually the results can (with some additional effort) be made accessible to a broad audience.
- Since the method explicitly discusses and incorporates subjective or "framing" issues, it is at least open and transparent to users. The method is inherently dynamic – that is what it does best
- The results can be aggregated to get a single benefits number as needed.
- Participants in the mediated modeling process gain deep understanding of the process and products. Those who have not participated can easily view and understand the results if they invest the effort. Usually the results can (with some additional effort) be made accessible to a broad audience.
- Since the method explicitly discusses and incorporates subjective or "framing" issues, it is at least open and transparent to users.

Status as a method. As mentioned above, mediated models can contain explicit valuation components. In fact, if the goal of the modeling exercise is to consider trade-offs, then valuation of some kind becomes an essential ingredient. How these trade-offs and valuations get incorporated into the model, varies, of course, from exercise to exercise. Perhaps the best way to describe this process is with an example. The South African fynbos ecological economic model described by Higgins et al. (1997) is an illustrative example.

The area of study for this example was the Cape Floristic Region—one of the world's

smallest and, for its size, richest floral kingdoms. This tiny area, occupying a mere 90,000 km2, supports 8,500 plant species of which 68% are endemic, 193 endemic genera and six endemic families (Bond and Goldblatt 1984). Because of the many threats to this region's spectacular flora, it has earned the distinction of being the world's "hottest" hot-spot of biodiversity (Myers 1990).

The predominant vegetation in the Cape Floristic Region is fynbos, a hard-leafed and fire-prone shrubland which grows on the highly infertile soils associated with the ancient, quartzitic mountains (mountain fynbos) and the wind-blown sands of the coastal margin (lowland fynbos) (Cowling 1992). Owing to the prevalent climate of cool, wet winters and warm, dry summers, fynbos is superficially similar to California chaparral and other Mediterranean climate shrublands of the world (Hobbs, Richardson, and Davis 1995). Fynbos landscapes are extremely rich in plant species (the Cape Peninsula has 2,554 species in 470 km2) and plant species endemism ranks amongst the highest in the world (Cowling 1992).

In order to adequately manage these ecosystems several questions had to be answered, including, what services do these species-rich fynbos ecosystems provide and what is their value to society? A two-week workshop was held at the University of Cape Town (UCT) with a group of faculty and students from different disciplines along with parks managers, business people, and environmentalists. The primary goal of the workshop was to produce a series of consensus-based research papers which critically assessed the practical and theoretical issues surrounding ecosystem valuation as well as assessing the value of services derived by local and regional communities from fynbos systems.

To achieve these goals, an 'atelier' (or combined workshop/short course) approach was used to form multidisciplinary, multicultural teams, breaking down the traditional hierarchical approach to problem-solving. Open space (Rao 1994) techniques were used to identify critical questions and allow participants to form working groups to tackle those questions. Open space meetings are loosely-organized affairs which give all participants an opportunity to raise issues and participate in finding solutions.

The working groups of this workshop met several times during the first week of the course and almost continuously during the second week. The groups convened together periodically to hear updates of group projects and to offer feedback to other groups. Some

group members floated to other groups at times to offer specific knowledge or technical advice.

Despite some initial misgivings on the part of the group, the structure of the course was remarkably successful, and by the end of the two weeks, seven working groups had worked feverishly to draft papers. These papers were eventually published as a special issue of *Ecological Economics* (Cowling and Costanza 1997). One group focused on producing an initial scoping (or mediated) model of the fynbos. This modeling group produced perhaps the most developed and implementable product from the workshop: a general dynamic model integrating ecological and economic processes in fynbos ecosystems (Higgins et al. 1997). The model was developed in STELLA and designed to assess potential values of ecosystem services given ecosystem controls, management options, and feedbacks within and between the ecosystem and human sectors. The model helped to address questions about how the ecosystem services provided by the fynbos ecosystem at both a local and international scale are influenced by alien invasion and management strategies. The model consists of five interactive sub-models: a) hydrology; b) fire; c) plants; d) management; and (e) economic valuation. Parameter estimates for each sub-model were either derived from the published literature or established by workshop participants and consultants (they are described in detail in Higgins et al. 1997). The plant sub-model included both native and alien plants. Simulation of the model produced a realistic description of alien plant invasions and their impacts on river flow and runoff.

This model drew in part on the findings of the other working groups, and incorporates a broad range of research by workshop participants. Benefits and costs of management scenarios were addressed by estimating values for harvested products, tourism, water yield and biodiversity. Costs included direct management costs and indirect costs. The model showed that the ecosystem services derived from the Western Cape mountains are far more valuable when vegetated by fynbos than by alien trees (a result consistent with other studies in North America and the Canary Islands). The difference in water production alone was sufficient to favor spending significant amounts of money to maintain fynbos in mountain catchments.

The model was designed to be user-friendly and interactive, allowing the user to set such features as area of alien clearing, fire management strategy, levels of wildflower

harvesting, and park visitation rates. The model has proven to be a valuable tool in demonstrating to decision makers the benefits of investing now in tackling the alien plant problem, since delays have serious cost implications. Parks managers have implemented many of the recommendations flowing from the model.

There are several other case studies in the literature of various applications of mediated modeling to environmental decision-making, including valuation. Van den Belt (2004) is the best recent summary and synthesis. Some additional examples of mediated modeling projects where ecosystem service values were integrated are:

- Participatory Energy Planning in Vermont, Department of Public Service in Vermont, http://www.publicservice.vermont.gov/planning/mediatedmodeling.html
- Mediated Modeling of the impacts of Enhanced UV-B Radiation on Ecosystem Services (van den Belt et al, 2006)
- Ria Formosa Coastal Wetlands, (a case study in van den Belt, 2004)
- Upper Fox River Basin, (a case study in van den Belt, 2004)
- A consensus-based simulation model for management of the Patagonian coastal zone, (van den Belt et al. 1998)

Models can be downloaded from: www.mediated-modeling.com

Strengths/Limitations. Resources needed to implement the method vary from application to application. The method can deal with a broad range of available data and resources, probably better that most other methods, since the model can adapt to the resources available across different levels of data, detail, scope and complexity. As a rule of thumb, one can produce a credible mediated model in 30-40 hours of workshops; about 300-400 hours of organizing/modeling. Cost: about $40,000 - $100,000 depending on side activities.

The most serious obstacle seems to be the fact that this method is very different from the top-down approach most frequently used in government. It requires that consensus building be put at the center of the process, which can be very scary for institutions accustomed to controlling the outcome of decision processes. An institutonal mandate is important, however, to motivate various stakeholders to volunteer their time, knowledge and

energy to a mediated modeling process. The final outcome of this process cannot be predetermined.

The most serious obstacle seems to be the fact that this method is very different from the top-down approach most frequently used in government. It requires that consensus building be put at the center of the process, which can be very scary for institutions accustomed to controlling the outcome of decision processes. The final outcome of this process cannot be predetermined.

Treatment of Uncertainty. In terms of uncertainty, there are all the usual sources, but the difference is that the stakeholders are exposed to these sources as they go, and learn to understand and accommodate them as part of the process. The method is compatible with formal or informal characterizing of uncertainty, producing probability distributions in addition to point estimates.

Research needs. No research has yet been done on whether application of the process to exactly the same problem by multiple independent groups would yield "consistent and invariant" results. One would expect general consistency, but some variation between applications. This is an area for further research.

To evaluate the impact of a mediated modeling process, surveys have been used before and after a process in the past and this research would deepen the understanding about exactly what elements of a mediated modeling process contribute to the success of failure of these processes.

Key References

Bond, P. and Goldblatt , P. 1984. Plants of the Cape Flora. Journal of South African Botany Suppl. 13:1-455.

Checkland, P. 1989. Soft Systems Methodology, in J. Rosenhead (ed.) Rational Analysis for a Problematic World, John Wiley and Sons, Chichester, England.

Costanza, R. 1987. Simulation Modeling on the Macintosh Using STELLA, BioScience, Vol. 37, pp. 129 - 132.

Costanza, R., F. H. Sklar, and M. L. White. 1990. Modeling Coastal Landscape Dynamics. BioScience 40:91-107

Costanza, R. and M. Ruth. 1998. Using dynamic modeling to scope environmental problems and build consensus. Environmental Management 22:183-195.

Costanza, R., A. Voinov, R. Boumans, T. Maxwell, F. Villa, L. Wainger, and H. Voinov. 2002. Integrated ecological economic modeling of the Patuxent River watershed, Maryland. Ecological Monographs 72:203-231.

Cowling, R.M. (ed.). 1992. The ecology of fynbos. Nutrients, fire and diversity. Oxford University Press, Cape Town.

Cowling, R. and R. Costanza (eds). 1997. Valuation and Management of Fynbos Ecosystems. Special section of Ecological Economics vol 22, pp 103-155.

Ford, A. 1999. Modeling the Environment: An Introduction to System Dynamics Models of Environmental Systems. Island Press, Washington, DC

Gunderson, L. C. S. Holling, and S. Light (eds). 1995. Barriers and bridges to the renewal of ecosystems and institutions. Columbia University Press, New York. 593 pp.

Hannon, B. and M. Ruth. 1994. Dynamic Modeling, Springer-Verlag, New York.

Hannon, B. and M. Ruth. 1997. Modeling Dynamic Biological Systems, Springer-Verlag, New York.

Higgins, S. I., J. K. Turpie, R. Costanza, R. M. Cowling, D. C. le Maitre, C. Marais, and G. Midgley. 1997. An ecological economic simulation model of mountain fynbos ecosystems: dynamics, valuation, and management. Ecological Economics 22:155-169.

Hobbs, R.J., Richardson, D.M. and Davis, G.W. 1995. Mediterranean-type ecosystems: opportunities and constraints for studying the function of biodiversity. In: Mediterranean-type ecosystems. The function of biodiversity. G.W. Davis and D.M. Richardson (eds), pp 1-42. Springer, Berlin.

Kahnemann, D. and A. Tversky. 1974. Judgment Under Uncertainty, Science, Vol. 185, pp. 1124 - 1131.

Kahnemann, D., P. Slovic, and A. Tversky. 1982. Judgment Under Uncertainty: Heuristics and Biasis, Cambridge University Press, Cambridge.

Lyneis, J.M. 1980. Corporate Planning and Policy Design: A System Dynamics Approach, Pugh-Roberts Associates, Cambridge, Massachusetts.

Morecroft, J.D.W. 1994. Executive Knowledge, Models, and Learning, in J.D.W. Morecroft, and J.D. Sterman (eds.) Modeling for Learning Organizations, Productivity Press, Portland, Oregon, pp. 3 - 28.

Morecroft, J.D.W., D.C. Lane and P.S. Viita. 1991. Modelling Growth Strategy in a Biotechnology Startup Firm, System Dynamics Review, No. 7, pp. 93-116.

Morecroft, and J.D. Sterman (eds.) 1994. Modeling for Learning Organizations, Productivity Press, Portland, Oregon

Myers, N. 1990. The biodiversity challenge: expanded hot-spots analysis. The Environmentalist 10: 243-255.

Peterson, S. 1994. Software for Model Building and Simulation: An Illustration of Design Philosophy, in J.D.W. Morecroft, and J.D. Sterman (eds.) Modeling for Learning Organizations, Productivity Press, Portland, Oregon, pp. 291 - 300.

Phillips, L.D. 1990. Decision Analysis for Group Decision Support, in C. Eden and J. Radford (eds.) Tackling Strategic Problems: The Role of Group Decision Support, Sage Publishers, London.

Rao, S.S. 1994. Welcome to open space. Training (April): 52-55.

Richmond, B. and S. Peterson. 1994. STELLA II Documentation, High Performance Systems, Inc., Hanover, New Hampshire.

Roberts, E.B. 1978. Managerial Applications of System Dynamics, Productivity Press, Portland, Oregon.

Rosenhead, J. (ed.) 1989. Rational Analysis of a Problematic World, John Wiley and Sons, Chichester, England.

Senge, P.M. 1990. The Fifth Discipline, Doubleday, New York.

Simon, H.A. 1956. Administrative Behavior, Wiley and Sons, New York.

Simon, H.A. 1979. Rational Decision-Making in Business Organizations, American Economic Review, Vol. 69, pp. 493 - 513.

Van den Belt, M. 2004. Mediated Modeling: a systems dynamics approach to environmental consensus building. Island Press, Washington, DC.

Van den Belt, Marjan, Oscar Bianciotto, Robert Costanza, Serge Demers, Susana Diaz, Gustavo Ferryra, Evamaria Koch, Fernando Momo, Maria Vernet. 2006. Mediated Modeling of the impacts of Enhanced UV-B Radiation on Ecosystem Services. Photochemistry and Photobiology, 82: 865-877

Van den Belt, M.J., L.Deutch, Å.Jansson, 1998. A consensus-based simulation model for management in the Patagonia coastal zone, Ecological Modeling, 110:79-103Vennix,

J. A. M. 1996. Group Model Building : Facilitating Team Learning Using System Dynamics. Wiley, NY.

Vennix, J.A.M. and J.W. Gubbels. 1994. Knowledge Elicitation in Conceptual Model Building: A Case Study in Modeling a Regional Dutch Health Care System, in J.D.W. Morecroft, and J.D. Sterman (eds.) Modeling for Learning Organizations, Productivity Press, Portland, Oregon, pp. 121 - 146.

Westenholme, E.F. 1990. System Inquiry: A System Dynamics Approach, John Wiley and Sons, Chichester, England.

Westenholme, E.F. 1994. A Systematic Approach to Model Creation, in J.D.W. Morecroft, and J.D. Sterman (eds.) Modeling for Learning Organizations, Productivity Press, Portland, Oregon, pp. 175 - 194.

6. PUBLIC AND GROUP EXPRESSIONS OF SOCIAL/CIVIC VALUE

6.1. Introduction

Traditional economic valuation methods attempt to measure and aggregate the values that individuals place on changes in ecological systems and services based on their personal preferences as consumers of those systems and services. An alternative approach is to try to measure the values that groups of individuals place on changes in such systems and services explicitly in their role as citizens – social/civic valuation. This approach measures the monetary value that groups place on changes in the systems and services when asked to evaluate how much *the public as a whole* should pay for increases in such systems and services (public willingness to pay) or should accept in compensation for reductions in the systems and services (public willingness to accept). The value measurement purposefully seeks to assess the full "public regardedness" value, if any, that the group attaches to any increase in community well-being attributable to changes in the relevant systems and services.

Social/civic values, like values based on personal preferences, can be measured either through revealed behavior or through stated valuations. One principal source of revealed values for changes in ecological systems and services are votes on public referendum and initiative involving environmental decisions. Other public decisions, however, also may provide measures of social/civil values, including official community decisions to accept compensation for permitting environmental damage, and jury awards in cases involving damage to natural resources. Where revealed values are difficult or impossible to obtain, social/civil values also can be measured by asking "citizen valuation juries" or other representative groups the value that they, as citizens, place on changes in particular ecological systems or services.

6.2. Referenda and Initiatives

Brief description of the method: Referendum votes and other formal public decisions provide the basis for a set of valuation approaches that can yield monetized values, but use somewhat different logic than that of the conventional individually based revealed-preference and stated-preference methods. The outcomes of referenda (measures placed on the ballot by a legislative body), initiatives (ballot measures proposed by citizens), or other official public decisions directly express what the body politic as a collectivity values in terms of policy outcomes. These expressions may or may not correspond closely to the aggregated values of the individuals in the community in terms of outcomes. Referenda approaches (not to be confused with the "referendum format" often used for posing questions to solicit contingent valuation responses) provide information about the policy preferences of the median voter; under certain circumstances this information can tell us about the median voter's valuation of specific environmental amenities, and can even provide information, albeit weaker, about mean valuations of those who participate in the voting process.

There are four variants for analyzing referenda and initiatives:

- Referendum/initiative analysis
- Analysis of public decisions to accept pollution or resource depletion
- Referendum/initiative analysis followed by a survey
- Analysis of public decisions to accept pollution or resource depletion followed by a survey

Direct referendum/initiative analysis, with or without a follow-up survey, can evaluate tradeoffs between community and/or household costs (higher taxes, possibly job losses) and eco-system improvements (establishment or improvement of air, water, biodiversity protection, etc.). Direct analysis of public decisions to accept pollution or resource depletion, with or without a survey, can evaluate tradeoffs between community and/or household benefits (increase in tax base, job creation, infrastructure improvements, etc.) and eco-system deterioration (greater pollution, amenity reductions).

Text Box 13: Direct Analysis of Public Decisions to Accept Pollution or Resource Depletion

While the approach outlined above provides information about willingness to pay, there are some public decisions that can provide inferences for willingness-to-accept

decisions. These decisions involve a community's vote as to whether to permit the entry of a new firm or a new (or increased) economic activity despite the expectation that such permission will degrade the eco-system. The payment represents the ceiling on the community's valuation of the environmental amenities that are being relinquished. It is a ceiling because of the possibility that the community would have accepted a lower level of compensation, and if the community valued the forgone eco-system services more than the compensation, then presumably it would not have accepted the compensation. However, if there is a vote and the outcome is close, the calculated valuation can be considered to be close to the community's valuation.

The estimation task involves assessing the amount of environmental damage in physical terms and the amount of compensation in monetary terms. Typically this compensation will come in the form of additional sources of taxes, the value of infrastructure that the new entrants provide for the community, additional income earned by community members, etc. The per-household as well as per-community compensation would be relevant. For example, the entry of an air-polluting factory may be accepted only after the factory's owner commits to a certain number of jobs for the community, building a park, upgrading roads, contributing to the community's vocational program.

Obviously many "community decisions" to permit the entry of polluters or other activities that degrade the ecosystem are not amenable to this approach, because community leaders negotiate the level of benefits that the community will receive without a vote being taken, or the benefits or costs are difficult to estimate.

Text Box 14: Referendum/Initiative Analysis Followed by a Survey

Another variant that relies on referendum and initiative outcomes to make willingness-to-pay estimates consists of combining the voting outcome with a follow-up survey to determine the perceptions of the stakeholders. This variant amounts to a hybrid of the first variant and the "referendum format" contingent valuation approach. The floor of the willingness-to-pay value of the proposed eco-system improvements is estimated by determining the voters' perceptions of the eco-system improvements and costs proposed by a recent referendum or initiative. The respondents are asked whether they voted, how they

voted, and what they believed the benefits and costs of the proposal were. The analysis of results of the referendum/initiative are the same as direct analysis without a survey.

If, in addition to asking how respondents voted and their perceptions of the benefits and costs of the proposal, the randomly-sampled respondents who opposed the proposal are asked what (lower) cost would have induced them to vote for the proposal, and those who supported the proposal are asked how much more they would have been willing to pay, this approach also permits an estimate of aggregate and mean values, just as a standard contingent valuation study would, with less potential distortion arising from respondents' desire to be regarded in a favorable light. Thus the survey following a referendum or initiative can provide an internal cross-check of how much correspondence there is between the stated-preference approaches and the referendum or initiative findings.

It should be noted that in focusing on the benefits and costs that respondents report, rather than the actual benefits and costs that the referendum or initiative proposal specifies, the results do not reflect the community's formal decision. This is a significant difference in the philosophy underlying the standing of the results. That is, the first variant, even if it does not necessarily reflect the values that voters perceive, does represent what the voters have chosen. Different logics underlie their standing.

Text Box 15: Public Decisions to Accept Pollution or Resource Depletion Followed by a Survey

Just as the analysis of referendum and initiative outcomes can be augmented by determining voters' perceptions of the stakes, the ceiling of the willingness-to-accept value of eco-system deterioration can be estimated by determining the benefits perceived by voters who supported the arrangement accepting the entry of a polluting or depleting operation into the community, and their perceptions of the damage that would be done. Like the direct analysis of willingness-to-accept votes, if the arrangement was approved by the electorate, and the property rights are clear and transactions are low, the ratio of the perceived costs and compensation represents the ceiling of the median voter's valuation. The survey, best administered as soon as possible after the actual vote, would reveal what the community members interpreted the benefits and costs to be, thus bringing the valuation closer to individual values; but again, with the tradeoff that the results would not have standing as the "community's choice." If the survey includes the questions of the conventional contingent

valuation regarding how much each respondent would have been willing to accept, then the results would be even more robust in finding mean and aggregate valuations as well as median valuations.

How the method could be used as part of the CVPESS expanded & integrated framework. The method relates to steps 2 and 4 of the C-VPESS framework. These public-decision approaches do provide monetized values—of the community's formal decision and values, ceilings, or floors of the median voter's valuation. Because the approaches focus on the content of proposals before the voting public, they do not directly identify ecosystem service impacts as a natural scientist or engineer would, but they will reflect voters' assessments of ecosystem service impacts. The approaches focusing exclusively on the decision outcomes do not directly identify changes in ecosystems and ecosystem services that are of greatest concern to people, although the survey variants can include questions to elicit this information. The approaches do address ecological impacts that other monetized approaches may underestimate, in that participation in citizenship, in contrast with the private-utility decisions reflected in the standard revealed-preferences approaches, can reflect concern for community well-being ("public regardedness") insofar as voters hold such regard. The approaches do not involve inter-disciplinary collaboration among physical/biological and social scientists or ecologists. There is a very strong potential that a data bank of inferred values from fairly large numbers of referenda and initiatives would assist EPA in presenting ranges of value for benefit transfers.

Status as a method: The logic of using formal public outcomes to infer how much "society values" particular outcomes has been used primarily in the literature on health and safety. For example, the value of a "statistical life" has been estimated by calculating how much public policies commit to spend in order to reduce mortality rates from health or safety risks, or, conversely, how much economic gain is associated with public decisions that reduce safety (e.g., by examining official decisions of U.S. states to raise or lower speed limits, Ashenfelter & Greenstone [2004] estimated the market value of the time saved by getting to the destination more quickly, and from that estimated the value of the additional expected traffic fatalities). The logic of making valuation inferences from referenda and

initiatives has been addressed in a few publications, most directly in Deacon & Shapiro, 1975; and Shabman & Stephenson, 1996.

In comparing the valuations yielded by stated-preference approaches with those derived from public decisions, the studies typically show the inferences from public decisions to yield lower values—not surprising in light of the absence of the hypothetical element in the public-decision results. Although systematic comparisons with conventional revealed preference approaches are lacking, it is likely that the valuations of eco-system components calculated from public decisions would be higher, because public decisions do capture whatever elements of public-regardedness are present among the voters. The valuations based on public decisions have intrinsic validity within the paradigm that gives standing to the community votes as reflecting the policies that the public prefers.

Strengths/Limitations: Willingness-to-pay (WTP): The results will be most easily interpreted if the initiatives or referenda are: a) as focused as possible on a single dimension of environmental protection or amenity; b) free of ideological debate; c) confined to easily identifiable government costs rather than diffused and uncertain costs such as job losses.

Willingness-to-accept (WTA): Assuming that a) the vote is explicit; b) the expected damage is well specified, c) property rights are clearly held by the community (i.e., the community has the right to refuse entry), d) the community's gains can be easily estimated, and e) the transactions costs are low,

The most useful referenda or initiatives would propose direct costs to the voters, typically in the form of taxes, fees, or bonds to finance actions designed to improve or protect eco-systems. Referenda or initiatives that entail restrictions on development (such as more stringent emissions or effluent standards) are less useful, because of the uncertainty of the level and incidence of the economic impacts. Similarly, in order to isolate the values attributed to particular ecosystem benefits, referenda and initiatives that address only one objective, such as preserving habitats or reducing air pollution. With multiple objectives, the analysis cannot assign the willingness to pay to each component. Similarly, if it is clear that a referendum or initiative entails additional partisan political stakes (e.g., if it is widely viewed as a political test of a government official), the results are less illuminating in terms of the ecosystem values that the voters hold.

All of these approaches would work best when:

a) applied to the same jurisdiction (e.g., if Portland is considering another storm control issue, the analysis of the Portland referendum would be most appropriate), but can still be used via benefits transfer;
b) a unitary conservation or environmental benefit is involved;
c) the initiative or referendum outcome was a close vote (this yields stronger inferences about the actual valuation, rather than floors or ceilings);
d) extraneous issues (such as whether the vote is a "political test" on particular politicians, or the mode of financing is controversial) are unimportant;
e) surveys can be accomplished soon after the actual vote.

All of these approaches attempt to measure the sum total of values of improving or protecting eco-systems and eco-system services; therefore both means and ends (instrumental and intrinsic) values can be involved. All variants in principle could measure the values attributed to all types of services, expressed in terms of monetary values per unit of eco-system improvement or protection. The variants are flexible in terms of levels of data, detail and scope, inasmuch as initiatives, referenda and other public decisions have been made at all sub-national levels. The valuations can be aggregated across benefits and with other methods, as long as the scale and magnitude of benefits are roughly the same. While highly complex initiatives, referenda, and other public decisions are not good candidates for estimating value, the valuations generated from simpler cases can be used as inputs for complex applications.

Any EPA decision context calling for monetized valuation could employ any of these variants, either singly or as cross-checks with conventional revealed preference or sated preference approaches. Benefit transfer applications will be limited to cases of similar magnitudes of benefits, because of the likelihood that community decisions are highly sensitive to such magnitudes.

In uses that apply valuations directly to the jurisdiction previously experiencing the initiative, referendum or negotiation, the scale would be the same municipality, country or

state. For benefits transfer, the scale should also be the same, given the need for similar magnitude of benefits and costs mentioned above.

Making valuation estimates directly from referendum or initiative outcomes has two advantages over conventional valuation methods. Unlike the standard revealed-preference approaches, such as hedonic pricing or the travel-cost method, voting on referenda or initiatives will reflect as much (or as little) public-regardedness as the voters actually hold toward the objectives involved. Standard revealed-preference approaches reflect the private-utility-maximizing decisions of individuals who purchase homes, spend money to visit parks, etc.; these decisions do not reflect what individuals want for their communities. Voting affirmatively for referendum- or initiative-proposed public expenditures do elicit valuing on behalf of the community, insofar as the voters are so disposed. Of course, a voter may vote for or against a referendum or initiative proposal strictly out of concerns for herself and/or her family, but the outcome does not exclude the existence value component if it exists.

Unlike the conventional stated preference approaches such as contingent valuation, the analysis based on referendum or initiative outcomes is not subject to the possible distortions of hypothetically-posed choices. If a voter supports the referendum or initiative proposal, the vote contributes to the likelihood that the expenditures will actually occur and the costs will actually be borne. Some might argue that the chance that any one vote will decide the outcome of the referendum or initiative is remote, and therefore the vote is more of a symbolic act than a tradeoff choice. However, there are two important responses to this point. First, whatever the mix of motives of the voters, the outcome is the community's decision, and therefore has standing in and of itself. This is the same logic by which we accept elected officials as legitimate even if we are dubious about the motives or rationality of the voters. Second, even if a voter believes that the chances that his or her vote will make the difference are negligible, the vote is still an expression of support or opposition to the proposal. There is little reason to believe that a "yes" vote would reflect just the gratification of voting "yes" (especially in secret balloting) rather than a belief that the proposal merits support.

Another concern that some would level against inferences based on referenda or initiatives is that these votes are often subject to intense efforts by interest groups, advocacy groups, and even governments to manipulate public perceptions. This concern has two

aspects: whether the information on which voters base their decisions has been distorted, and whether the votes are swayed by appeals on one side or the other. The first aspect is more compelling: we certainly would be less willing to accept the validity of an estimate derived from voting decisions driven by serious misconceptions of the proposed benefits and/or costs. The outcome is still the official decision of that community, but the justification for using the result as the basis of benefits transfer to other communities would be very weak. On the other hand, the fact that referenda and initiatives are often subject to intensive campaigns of persuasion may be considered a virtue rather than a drawback, insofar as it would provide more information on both sides. In addition, the fact that individuals are exposed to efforts at persuasion is by no means confined to referenda and initiative contests: respondents to contingent valuation surveys have of course been subjected to many years of promotional activities by environmental groups; people who travel farther to a particularly popular national park such as Yosemite have been influenced by all sorts of communications extolling its virtues. In short, efforts at value persuasion are pervasive, and in any event should not be a basis for rejecting the significance of decisions of individuals exposed to those efforts. The philosophical basis underlying the use of referenda or initiatives, namely that the public's preferences are legitimately shaped by the political process, and that the public's policy preferences are important beyond how the public values the outcomes that these policies may produce, is quite different from the so-called "progressivist" position that individuals' values should be determined in isolation of "politics" (Sagoff 2004: 177-178).

Another difference in philosophical basis is that the referendum and initiative results reflect intensity of attention to the issue, at least insofar as those who do not care enough to vote are excluded from the analysis. From the progressivist, technocratic perspective, everyone's values ought to be incorporated, because the policies ought to maximize utility (i.e., the consequences of public decisions) regardless of whether specific individuals are mobilized to take action. On the other hand, prominent strains of pluralist democratic theory regard intensity as a fully legitimate factor in determining policy outcomes (Lowi 1964).

One limitation of estimating values from referendum or initiative outcomes is that it is sometimes difficult for voters to assess the actual stakes involved. The benefits will often have to be predicted (e.g., how much biodiversity will a reserve really safeguard; how much flooding will the flood-control system actually prevent?), entailing a certain amount of

uncertainty. The benefits that do occur will often be community-wide, with some uncertainty as to how much an individual or particular household can take advantage of the benefits. On the cost side, the burden of a tax increase or bond measure on household expenditures may be very difficult for the typical voter to estimate, and the impacts of development restrictions may be even more difficult in light of the uncertainty as to which families would ultimately be affected. Insofar as the costs specified by the referendum or initiative are not easily translatable into household budget terms, the outcome, though it is still "the community's decision," is less revealing about the values held by the voters.

The outputs of these approaches should be easy to understand and to communicate to the public. It is a significant advantage to be able to say that the valuation of an eco-system component has been estimated on the basis of how communities have decided what these components are worth.

Text Box 16: Referenda and Initiatives Used to Validate Contingent Valuation

In addition to taking the valuation derived from the analysis of public decisions as an input in itself, the analysis of public decisions, particularly referenda and initiatives, can be used to validate the results of other valuation methods. Several studies have compiled the results of initiatives and/or referenda in order to try to validate more conventional valuation techniques, especially contingent valuation (Kahn & Matsusaka (1997), List & Shogren (2001; 2002), Murphy et al. (2003), Polasky, Gainutdinova & Kerkvliet. (1996), Schläpfer, Roschewitz, & Hanley (2004). Vossler & Kerkvliet (2003)). As Arrow et al. (1993) recommend:

The referendum format offers one further advantage for CV. As we have argued, external validation of elicited lost passive use values is usually impossible. There are however real-life referenda. Some of them, at least, are decisions to purchase specific public goods with defined payment mechanisms, e.g., an increase in property taxes. *The analogy with willingness to pay for avoidance or repair of environmental damage is far from perfect but close enough that the ability of CV-like studies to predict the outcomes of real-world referenda would be useful evidence on the validity of the CV method in*

> *general. The test we envision is not an election poll of the usual type. Instead, using the referendum format and providing the usual information to the respondents, a study should ask whether they are willing to pay the average amount implied by the actual referendum. The outcome of the CV-like study should be compared with that of the actual referendum. The Panel thinks that studies of this kind should be pursued as a method of validating and perhaps even calibrating applications of the CV method.*(emphasis added)

Does this method incorporate any specific ways of treating uncertainty? Is there any approach unique to this method? There are two distinct sources of uncertainty involved with this approach, depending on which variant is employed and how the outcomes are interpreted. If the referendum or initiative results are used without a follow-up survey, and the results are interpreted as indicating the aggregation of individual valuations, then there is uncertainty as to whether the voters understood the benefits and the payments accurately. If the results are interpreted as the community's preference per se, then the result is accurate in itself, as long as vote miscounting is not an issue.

The follow-up survey provides a way to determine whether voters understood the benefits and payments accurately. However, like any survey it also has its own sources of uncertainty: biases in which voters agree to respond to the survey, and untruthfulness in the individual responses. An additional source of potential uncertainty would arise if individuals who did not vote respond to the survey because of error on the part of the survey team. Despite these potential pitfalls, the follow-up survey (equivalent to a contingent valuation study) would serve as a cross-check on the referendum or initiative results.

Another source of uncertainty in undertaking a benefits transfer of valuation based on referenda or initiatives is that communities where these efforts are tried may be atypical; for example, it is possible that referenda and initiatives are more likely to be launched in communities with a stronger commitment to conservation. However, if enough straightforward referenda and initiatives are analyzed and put into comparable terms, including those that failed to pass, the range of results would provide more robust information than any single result.

Research needs: The research needed to make the results of public decisions through referenda and initiatives would consist of the creation of a data bank of referenda and initiative outcomes, optimally screening out those involving multiple, confounding elements. The analysis of the valuation of benefits or damage would be straightforward calculation of the ratios of benefits or costs to the per-household costs, when such ratios can be deduced from simple referendum or initiative choices. The Resources for the Future efforts seem to be progressing in this direction. The survey variants would involve considerably more effort of developing the questionnaire, administering it immediately after a referendum or initiative, and analyzing the additional information, yet the results would provide information on both median and mean valuation. Once model surveys are developed, they could be used with minor adaptations in different settings. Three researchers could produce a credible data base,

Key References

Arrow, K., R. Solow, et al. (1993). Report of the NOAA Panel on Contingent Valuation. Washington, D.C, Government Printing Office.

Ashenfelter, O. and M. Greenstone (2004). "Using Mandated Speed Limits to Measure the Value of a Statistical Life." Journal of Political Economy 112: S226-S267.

Deacon, R. and P. Shapiro (1975). "Private preference for collective goods revealed through voting on referenda." American Economic Review 65: 793.

Kahn, M. E. and J. G. Matsusaka (1997). "Demand for environmental goods: Evidence from voting patterns on California initiatives." Journal of Law and Economics 40: 137-173.

List, J. and J. Shogren (2002). "Calibration of Willingness-to-Accept." Journal of Environmental Economics and Management 43: 219-233.

Schlapfer F, Roschewitz A and Hanley N (2004) " Validation of stated preferences for public goods: a comparison of contingent valuation survey response and voting behaviour." Ecological Economics, 51, 1-16.

Shabman, L. and K. Stephenson (1996). "Searching for the correct benefit estimate: Empirical evidence for an alternative perspective." Land Economics 72: 433-49.

Murphy, J. J., P. G. Allen, et al. (2003). A Meta-Analysis of Hypothetical Bias in Stated Preference Valuation, Department of Resource Economics University of Massachusetts Amherst, MA.

Vossler, C. A. and J. Kerkvliet (2003). "A criterion validity test of the contingent valuation method: Comparing hypothetical and actual voting behavior for a public referendum." Journal of Environmental Economics and Management 45(3): 631-49.

Schläpfer, Roschewitz, & Hanley (2004).

6.3. Citizen Valuation Juries

Description of the Method. Another potential process for attempting to measure the social/civil value of changes to ecological systems and services is to assemble and query a representative group of citizens (a "citizen valuation jury"). Like initiatives and referenda, citizen valuation juries provide information on social/civic values, but they measure stated rather than revealed value, and they incorporate elements of the "deliberative valuation" processes described in Section 6. Citizen valuations juries are also similar to contingent valuation surveys except that: a) juries are asked to determine how much the *public* should pay or accept in compensation for a specified ecological change (rather than being asked how much they would pay or accept as individuals); b) valuation juries are often asked to agree on a common value for the ecological change (rather than being asked for individual values that the expert then aggregates or otherwise combines); c) juries deliberate together as a group before determining value; and d) juries are provided with more extensive information about the ecological change and can be aided in their deliberations.

As discussed below, citizen valuation juries are a largely untried mechanism for determining the social/civic value of ecological changes. Although citizen juries have been used an aid to various governmental decisions, juries have generally been asked to rank governmental options rather than to determine monetary values. Citizen juries, however, can also be asked to determine either a social/civic willingness to pay ("public WTP") or a social/civic willingness to accept ("public WTA") for any particular ecological change (Blamey et al., 2000). For public WTP values, citizen valuation juries can be asked to determine the highest levy, tax, or other form of payment that the government should pay to obtain a particular ecological benefit. For public WTA values, citizen valuation juries can be

asked to determine the highest monetary sum that the government should accept to avoid a particular ecological loss.

Although there is little experience with citizen *valuation* juries, citizen juries more generally have been used in a significant number of settings to inform various governmental choices (see box below). The process of forming and utilizing citizen juries has varied widely. In the typical situation, a small group of citizens, typically ranging from a cross-section of 12 to 20 persons, have been drawn from the relevant population. Approaches have differed as to how best to choose the jurors. Given the small size of citizen juries, there is an inevitable tension between choosing jurors to reflect the demographic characteristics of the relevant population as a whole and choosing jurors that represent the interests of major stakeholders. Although larger juries would reduce some of the tensions involved in juror selection, larger juries are likely to find it more difficult to reach agreement within a realistic time frame. Most citizen juries to date have chosen using random sampling or stratified random sampling (Blamey et al., 2000).

Once a citizen jury is chosen, the jury then meets and deliberates over a multi-day period, during which it hears and questions expert witnesses, deliberates in small and large groups, and agrees on a final recommendation to the sponsoring governmental body. These group deliberations allow jurors to hear alternative perspectives, test ideas, and carefully work through the valuation exercise. Citizen juries use several different techniques to provide information to the jurors. In some cases, the government or an expert facilitator chooses what information to provide to jurors, while in other cases, relevant interest groups make individual presentations to the jury. Jurors also can be permitted to request information and pose questions directly to expert witnesses (Blamey et al. 2000). Two factors should guide choices among the processes for providing information to the jurors: a) ensuring that jurors have all information that they believe is valuable to their valuation exercise, and b) ensuring that the information is balanced and not biased toward any particular result.

Another important choice in designing a citizen jury is the process by which the jury will make decisions. In most cases, juries are asked to arrive at a group decision. Decision making rules in this context include simple majority votes of the jury, consensus (where a majority favors the valuation and no juror opposes it), and unanimous agreement.

In other cases, the jury is not asked to agree on a decision, but researchers instead ask for input from individual jurors that is then combined into an overall evaluation. Measures of central tendency (means or mediums of the valuations provided by the individual jurors) can be used to develop a valuation measure in this context.

Experiments indicate that citizen juries often produce significantly different valuation results from economic or socio-psychological surveys. The additional information available to jury members, the opportunity to spend time thinking about the appropriate valuation, and the stress on collective rather than individual values all appear to generate significant changes in valuation (Alvarez-Farizo & Hanley 2006). The jury's valuation of particular ecological improvements, however, can either increase or decrease compared to the results obtained through economic surveys (Alvarez-Farizo & Hanley 2006).

Because contingent valuation methodology and other traditional economic measurement approaches seek a very different valuation than citizen valuation juries, juries should not be seen as a substitute for the traditional approaches. Governmental agencies should employ citizen valuation juries as a supplement to and check on traditional economic valuation approaches. Decisions whether to pursue particular regulations or other governmental actions should consider estimates of both private and public value, along with the strengths and weaknesses of each approach.

Text Box 17: AValuation Exercise Illustrating Use of Citizen Juries

In one experiment, a citizen jury was used to examine the economic value of the control of a particular exotic weed, Bitou Bush (Chrysanthemoides monilifera L. Norl. ssp. rotundata) in an Australian national park (James & Blamey 2000). A jury of 14 was selected, using a two-phase telephone survey, in order to be representative of the New South Wales population on the basis of: gender; age; place of residence; rating of the environment in relation to other social issues; occupation; income; income source; and education. The jury met for three days during which they heard and questioned seven expert witnesses. Prior to the hearings, jurors received training in note taking and questioning of witnesses, in order to maximize their ability to use the information provided.

In one of the charges, the jury was given two options: (Option #1) the then current situation in which weeds were controlled on 3000 hectares per year, and (Option #4) an alternative management regime in which weed control would be expanded to 9600 hectares per year. The jury was then given the following charge: "How high would a park management levy have to be, before the jury would recommend Option 1 rather than Option 4 …? In other words, how high would the levy have to be before the … public would be no better off under Option 4 than Option 1?" The jury first decided that a progressive levy, calculated as a percent of gross income, was most appropriate. After discussing two proposed levies (0.1% and 0.25%) , the jury voted eight to two in favor of a levy of 0.1%. In a survey following the jury exercise, jurors reported that they found the valuation exercise to be both interesting and worthwhile.

Text Box 18: Other Valuation Exercises Using Citizen Juries

Citizen juries can be used in various forms of valuation exercises and need not produce monetary valuations. The major use of citizen juries to date, for example, has been to help governments rank options for achieving particular goals – e.g., reduction of traffic in an urban area (Kenyon et al. 2001). Information obtained during ranking deliberations can provide valuable insights for other valuation exercises (Aldred & Jacobs 2000). Citizen juries also have been combined with choice modeling to determine paired rankings of various ecological characteristics (Alvarez-Farizo & Hanley 2006).

Citizen juries also do not need to produce a collective value. In some experiments, for example, juries deliberate as a group, but members of the jury then report their valuations on an individual basis (Alvarez-Farizo & Hanley 2006).

Text Box 19: Other Group-Based Approaches to Valuation

Concerns whether contingent valuation surveys provide sufficient time and information for survey respondents to generate reliable estimates of the value of often complex ecological changes has led a number of researchers to investigate other group-based

approaches to valuation. Under the "Market Stall" ("MS") approach, for example, researchers meet with survey subjects in two one-hour meetings, separated by a week, and encourage the participants to discuss their valuations with household members and friends between the two sessions. Unlike citizen valuation juries, the MS approach asks survey subjects for their personal valuations, based on individual preferences and incomes, rather than social/civic valuation. Respondents are asked for their personal valuations in a confidential written survey at the end of the second meeting. In Macmillan et al. (2002), the WTP measures obtained through the MS approach were significantly lower than the WTP measures generated from CV interviews, which is consistent with other studies that show a decline in WTP when survey subjects are provided additional time to consider their answers (Whittington et al. 1992).

Relation of Method to the CVPESS Expanded and Integrated Framework.

Citizen juries are potentially useful both to identify socially important assessment endpoints and to attach a monetary value to changes in the assessment endpoints. Use of this method relates to steps 2 and 4 of the C-VPESS proposed valuation process.

Because citizen juries consist of representative members of the public, citizen juries also expand the role that the public plays in valuations of changes in ecological systems and services. Members of citizen juries actively evaluate information regarding changes, are permitted to ask questions of experts, and consciously deliberate over the appropriate social/civic value of the change.

Status as a Method. Only a few efforts have been made to date to use citizen juries to generate monetary estimates of the social/civic value of environmental changes. Citizen valuation juries should be considered experimental for the moment and should not be used to make significant governmental decisions until further research has been conducted on both the efficacy of the process and the appropriate jury processes. Given the potential use of citizen valuation juries to evaluate social/civic values, however, this is an area in which research can be valuably focused.

Strengths/Limitations. One of the major strengths of the citizen valuation jury is that, like referenda and initiatives, the citizen valuation jury incorporates public-regardedness.

Jurors are asked to provide a valuation based on the perceived impact of an ecological change on the entire community rather than on his or her individual preferences alone. Citizen valuation juries thus incorporate a broader concept of value than standard contingent valuation approaches and place the jurors in a position similar to that of the governmental decision makers who are being advised.

Citizen valuation juries avoid a number of potential concerns regarding referenda and initiatives as a source of social/civic valuation information. First, the jury process ensures that juries receive more information regarding the ecological change than most voters receive prior to voting on an initiative or referendum. Second, because the jury evaluation process can be carefully structured, citizen evaluation juries are less subject to undue influence from political interest groups than are votes on referenda and initiatives. Finally, there are a limited number of referenda and initiatives from which valuations can be derived, while citizen valuation juries can be asked to assess a valuation for any ecological change.

Citizen valuation juries build on a well-established legal institution in the United States – the criminal and civil jury system. The legal system uses juries to decide whether to initiate criminal prosecutions, determine guilt and innocence in criminal cases, decide between life and death in capital cases, and assess damages in often complex civil cases. Most adult members of the public have served as jurors, understand the importance of the role they assume, and act deliberately and responsibly.

Citizen valuation juries suffer from the hypothetical character of all stated-value methods of valuation. Because the juries do not themselves determine governmental policy, the juries may not reveal what they actually believe to be the social/civic value of an ecological change. The hypothetical character of jury valuations could be eliminated by providing that the valuations will directly determine whether particular governmental actions will be taken, but the government is unlikely to want to delegate its decision making powers to citizen juries. Despite concerns over hypothetical inquiries, experiments with citizen juries indicate that jurors approach their valuation task in a responsible fashion and reach well-thought-out conclusions (Aldred & Jacobs 2000).

Citizen juries also raise a number of other unique concerns. Some economists, for example, have worried that group dynamics and “norms” might reduce the reliability of jury decisions. Some jurors, for example, might not wish to be perceived as disagreeing with

others, while some jurors may be able to dominate the discussion and result. Feedback from jury experiments, however, has not indicated that such jury dynamics are significant (Macmillan et al. 2002). Trained facilitators may be able to overcome any structural pathology that might otherwise arise and should be involved in any valuation exercise involving citizen juries.

As discussed earlier, the choice of jurors also poses difficulties. Because of the small size of typical citizen juries, a demographic cross-section of the public may not adequately represent all interest groups. Choosing representatives of different interest groups to serve on citizen juries, however, may yield a jury that does not adequately represent demographics. Small citizen juries, moreover, will inevitably fail to fully represent the public as a whole. In order to ensure that jurors are other-regarded, experiments suggest that the government should choose a jury that is as demographically representative as possible (typically through stratified random sampling), so that the jury is at least symbolically representative, and then instruct the jury to adopt an impartial stance in its deliberations (Brown et al. 1995, Blamey et al. 2000).

Treatment of Uncertainty. Citizen valuation juries raise many of the same uncertainties as traditional methods of economic valuation. The small size of citizen juries, however, raises an additional uncertainty factor.

Research Needs. Because there is little experience with citizen valuation juries, further research is needed on a variety of topics before EPA should consider adopting the approach to develop social/civic valuations for decision making purposes. Key questions include:

- Do citizen valuation juries arrive at different valuations than individual respondents to CV surveys? If so, how and why do the valuations differ?
- How stable are valuations provided by citizen juries? How much variation exists among the valuations produced by different citizen juries?
- How do jury selection processes affect the valuations of the jury? What methods exist to overcome the inevitable bias arising from the small size of citizen juries?
- How should information be provided to citizen valuation juries? What are the advantages and disadvantages of highly structuring the information that is provided to a jury, versus permitting the jury to determine the information that it receives?

- How do decision making rules (e.g., consensus versus unanimity) affect valuations? What are relevant considerations in choosing among the different decision making rules?

Key References

Aldred, J. and M. Jacobs (2000). "Citizens and Wetlands: Evaluating the Ely Citizens' Jury." Ecological Economics **34:** 217-232.

Alvarez-Farizo, B. and N. Hanley (2006). "Improving the Process of Valuing Non-Market Benefits: Combining Citizens' Juries with Choice Modelling." Land Economics **82**(3)**:** 465-478.

Blamey, R.K., et al. (2000). Citizens' Juries and Environmental Value Assessment. Canberra, Australia, Research School of Social Sciences, Australian National University.

Brown, T.C., et al. (1995). "The Values Jury to Aid Natural Resource Decisions." Land Economics **71**(2): 250-260.

Gregory, R. and K. Wellman (2001). "Bringing Stakeholder Values into Environmental Policy Choices: A Community-Based Estuary Case Study." Ecological Economics **39:** 37-52.

Kenyon, W. and N. Hanley (2001). "Economic and Participatory Approaches to Environmental Evaluation." Glasgow, U.K., Economics Department, University of Glasgow.

Kenyon, W. and C. Nevin (2001). "The Use of Economic and Participatory Approaches to Assess Forest Development: A Case Study in the Ettrick Valley." Forest Policy and Economics **3**(1)**:** 69-80.

Macmillan, D.C., et al. (2002). "Valuing the Non-Market Benefits of Wild Goose Conservation: A Comparison of Interview and Group-Based Approaches." Ecological Economics **43:** 49-59.

McDaniels, T.L., et al. (2003). "Decision Structuring to Alleviate Embedding in Environmental Valuation." Ecological Economics **46:** 33-46.

O'Neill, J. and C.L. Spash (2000). "Appendix: Policy Research Brief: Conceptions of Value in Environmental Decision-Making." Environmental Values **9:** 521-536.

7. METHODS USING COST AS A PROXY FOR VALUE

7.1. Replacement Costs

Brief description of the method. The replacement cost method does not fall neatly into either revealed preference or stated preference. This method, also called avoided cost, uses the cost of replacing ecosystem services with a human engineered system as an estimate of the value of those ecosystem services. For example, an estimate of the value of conserving a watershed that naturally provides clean drinking water could be derived by estimating the cost of building a water filtration plant that would provide the same service. The replacement cost approach should not be confused with applications of "averting behavior" methods of estimating values which are based upon observed voluntary behavior.

Status as a method. Examples of use of replacement cost include estimates of the value of protecting watersheds for municipal drinking water supply, the most famous of which is the Catskills watersheds (see NRC chap 5 for references). Avoided cost of illness is an approach EPA has used successfully to account for certain human health benefits of environmental regulations.

Strengths/Limitations. Replacement cost can be a valid measure of economic value if three conditions are met: the human-engineered system must provide services of equivalent quality and magnitude, the human-engineered system must be the least costly alternative, and individuals in aggregate must be willing to incur these costs if the natural service were not available (Bockstael et al. 2000; Shabman and Batie 1978).

All valuation methods can be misconstrued, applied incorrectly and misinterpreted, however replacement costs require special caution. In the case of replacement costs, there are particular conditions, under which replacement cost can approximate the economic value of environmental protection.

There is great potential for abuse in using replacement costs to estimate the value of ecosystem services and it should be used with care. The loss of an ecosystem service does not necessarily mean that the public would be willing to pay for the least cost alternative. If so, use of replacement cost is invalid. Even when the benefits of the service exceed the least cost method of providing the service, replacement cost does not measure the willingness to pay for an environmental improvement or the avoidance of harm. It merely represents the value (avoided cost) of not having to provide the service via human engineered approaches.

Still, if there are alternative ways of producing the same service and that service would be demanded if provided at the least cost human engineered alternative way of providing the service, then replacement cost is a valid measure of the change in economic value from loss of the ecosystem service.

Key References

Bockstael, N. E., A. M. Freeman, et al. (2000). "On measuring economic values for nature." Environmental Science and Technology **34**: 1384-1389.

National Research Council (2004). Valuing Ecosystem Services; Toward Better Environmental Decision-Making. Washington, D.C., The National Academies Press.

Shabman, L. A. and S. S. Batie (1978). "The Economic Value of Coastal Wetlands: A Critique." Coastal Zone Management Journal **4**(3): 231-237.

7.2. Tradeable permits

In the case of tradable permits, there are no conditions under which the cost of permits could be used as a proxy for economic value.

Emissions permit trading has been allowed under the Clean Air Act since the 1990 Amendments. Under a cap-and-trade system, such as that used by EPA to reduce sulfur dioxide emissions, the regulatory body determines the total number of permits available and some means of allocating permits among regulated sources. A regulated source must ensure that it has sufficient permits to cover its activities or face penalties. In the example of tradable emissions permits, a regulated source can take actions to reduce its own emissions and/or purchase permits from other sources. For those firms with higher marginal cost of pollution control, cost savings can occur if they purchase emissions reduction credits from firms with lower pollution control costs. Similarly, firms with relatively low pollution control costs can profit by undertaking greater abatement and selling extra permits. In so doing, trading can reduce overall costs of compliance. Tradable permits schemes have been proposed in fisheries management in the form of individual transferable quotas (ITQs), and in land conservation in the form of transferable development rights (TDRs).

It has been suggested that the price of a tradable permit is a proxy for the economic value of provision of environmental quality or conservation. However, this confuses the notion of costs and benefits. In market equilibrium, the price of a tradable permit is equal to the marginal cost of supplying a unit of environmental quality or conservation covered by the permit. Permit price need not bear any relation to benefit of environmental quality or conservation. If there are a large number of permits issued relative to demand for permits then permit price will be low; with few permits, price will be high. This does not necessarily mean that the value of environmental quality or conservation is low (or high). Permit price only reflect value if price equals the marginal benefit of environmental improvement or conservation, which occurs only if the number of permits issued is such that marginal costs and marginal benefits equal. But issuing the right number of permits to get marginal cost equal to marginal benefits requires knowing marginal benefit in the first place. There is no way to be confident that tradable permit prices reflect value without already knowing value. In other words, tradable permit prices do not constitute a valuation methodology capable of generating information about values.

7.3. Habitat Equivalency Analysis

Brief description of the method. Habitat Equivalency Analysis (HEA) is a valuation method developed to calculate compensation for loss of ecological services resulting from injury to a natural resource over a specific interval of time. The method calculates the amount of habitat/resource to be created or enhanced to provide the same level of services over time as were lost due to the injury. The method simultaneously quantifies injury and scales the size of restoration. An implicit assumption is that the value of a unit of lost service and that of the replacement service are comparable (of the same type, quantity, and comparable value). The primary steps in performing a HEA are as follows:

a) Document and quantify injury
b) Identify and evaluate replacement project options
c) Scale the replacement project to compensate for injury
d) Determine the appropriate means of compensation
   - Cost selected restoration options

- Set performance standards for compensatory restoration projects

The HEA approach focuses on scaling value on a service-to-service basis (e.g., acres of service) and a similar methodology, Resource Equivalency Analysis, focuses on scaling value on a resource-to-resource basis (e.g. Number of Individuals in Listed species). Therefore in expressing values, HEA relies on biophysical units such as acres of habitat-service and services projected over time in service acre-years. Since this is a restoration/compensation method that is projected into the future, the final unit is a Net Present Value (NPV) measure of the service into the future stated as Discounted Service Acre Years (DSAYs).

How the method could be used as part of the CVPESS expanded & integrated framework. The spatial scale at which HEA has typically operated has been at the level of local to regional impacts. HEA also operates over a time scale in that it involves compensation for injury or estimates services produced by an action over time as well as allows for time for restoration projects to mature to full service value.

With regard to where to place HEA within the C-VPESS framework, it would seem to bridge a number the process elements. Because HEA simultaneously quantifies the injury and scales the value of restoration it would appear to be both a valuation and impact assessment tool. As well, it would seem to be flexible enough to work both as a tool for estimating non-monetary values under OPA and monetary valuation under CERCLA (the latter being accomplished by including the costs to restore a DSAY as the measure of value). Additionally, although HEA and REA are currently used in the post-hoc context of injury, damages and compensation, there is no reason that these methods are constrained to managing adverse outcomes after the fact. These methods could just as easily be used to compare alternative future actions to identify the action with the least impact and to compare alternative actions to identify which will yield the most service. These methods or variations could be a fruitful avenue for the Agency to explore through research and development activities.

The HEA methodology relies on structural or spatial measures of ecological components such as acres of habitat. Specific service types such as provisioning, regulating,

cultural and supporting services as expresses in the Millennium Ecosystem Assessment framework are not identified or expressed but would be considered to be present and operating. This is not problematic because of the simplifying assumption that the lost acre equals the replacement acre and therefore the service is present in both.

Temporal assumptions are very important in working with HEA especially in a damage assessment. Questions such as the following need to be answered or estimated:

- How long has the injury or lost service been in place?
- How much time is required to implement the restoration project?
- How long will the restoration project take before it reaches full replacement service?

Obviously the answers to these questions can have a significant impact on the estimated compensatory value required to offset the injury. In HEA a discount rate must be selected for the NPV calculations

There are some crucial assumptions associated with theHEA method. It can be used only when values per unit of replacement services and lost services are comparable, when it is possible to use a common metric to define an injury and the value of replacement services, and when replacement of ecological services is feasible and measureable.

<u>Status as a method,</u> The HEA approach was originally developed in 1992 to quantify damages associated with contaminated wetlands (CITATION?) and has since been applied to cover injuries due to chronic contamination, spills, and vessel groundings in a variety of habitats (CITATIONS?). HEA is currently used in Natural Resource Damages Assessment (NRDA) under Oil Pollution Action (OPA) And CERCLA (Superfund). The purpose of NRD actions is to make the public whole for injuries to natural resources that result from the release of hazardous substances or oil. It is important to note that restoration for damages is distinct from remediation activities. The method has also been tried by the Army Corps of Engineers for wetland permitting.

Interestingly under these two regulatory frameworks there is a different focus on compensation. Under Superfund actions, compensation for damages is focused on monetary compensation, which requires value be scaled to cost, while under OPA the focus is on

resource compensation. Under OPA the question is how many public resources the public requires to be made whole for their loss, so therefore value is scaled from resource to resource. As noted previously, there are no barriers to applying these methods or adaptations of them in proactive support of decisions. Therefore the Agency should explore such proactive applications of HEA and REA in other regulatory contexts and especially in collaborative partnerships with conservation as a focus.

Strengths/Limitations. The HEA method can be used as a way to value non-market services often overlooked by other valuation methods when the specific assumptions associated with HEA can be met. The method is not complicated mathematically. It is by nature inter-disciplinary because determination of comparability per unit of replacement services and lost services requires collaboration between ecologists and economists.

Since HEA and REA are currently applied to support regulatory actions which link to a litigation process, to define compensation the analysis and supporting data need to be legally defensible with regards to analytical quality. The chief analytical difficulty is to determine defensible input parameters, especially an appropriate metric for lost and restored services.

The HEA method is not appropriate for standard benefit-cost analysis, where the goal is to determine optimal (efficient) allocation of scarce resources. The cost of compensatory restoration projects should not be communicated as the benefit of the resources to the public.

Treatment of Uncertainty. To some degree, uncertainty is dampened (or ignored) by the principal operating assumption that the habitat lost and the habitat created are equal or eventually will be equal. In practice such an assumption is tracked by monitoring the performance of projects in the field to see if they do or don't yield adequate replacement service.

Research needs. This method will be strengthened by improved understanding of how to define metrics to most fully characterize injured and replacement ecological resources and services. Another area for research concerns comparability of ecological resources and guidelines for considering on-site, in-kind changes in resources and ecological services, as compared with off-site and out-of-kind resources. Improved understanding of the concepts of ecological capacity and the opportunity and payoff for restoration in the evaluation and design of restoration projects will also strengthen the method.

Key References

Habitat Equivalency Analysis: An Overview." Policy and Technical Paper Series, No. 95-1, (Revised 2000). NOAA. Washington, D.C

Damage Assessment and Restoration Plan and Environmental Assessment for the Point Comfort/Lavaca Bay NPL Site Recreational Fishing Service Losses, National Oceanic and Atmospheric Administration et. al 2001. Washington, D.C.

Chapman, D., N. Iadanza, and T. Penn. 1998. Calculating Resource Compensation: An Application of the Service-to-Service Approach to the Blackbird Mine Hazardous Waste Site. National Oceanic and Atmospheric Administration, Damage Assessment Center Technical Paper Series

NOAA Coastal Service Center. Habitat Equivalency Analysis - www.csa.noaa.gov/economics/habitatequ.htm

Dunford, R.W., T.C. Ginn and W.H. Desvousges. 2004. The Use of Habitat Equivalency Analysis in Natural Resource Damage Assessment. Ecological Economic. Volume 48: pp 49-70

APPENDIX A: SURVEY ISSUES FOR ECOLOGICAL VALUATION: CURRENT BEST PRACTICES AND RECOMMENDATIONS FOR RESEARCH

Background material prepared by Drs. Jon Krosnick and Allyson Holbrook for C-VPESS Consideration. Open issue: how to incorporate material from this into an appendix for the report.

Defining Survey Research

Survey research entails collecting data via a questionnaire from a sample of elements (e.g., adults who are American citizens living in the United States) systematically drawn from a well-defined population (e.g., all adult American citizens living in the United States; see Babbie, 1990; Fowler, 1988; Frey, 1989; Lavrakas, 1993; Weisberg, Krosnick, & Bowen, 1996). Conducting a survey involves three principal phases: (1) drawing a sample of respondents to represent a population, (2) collecting data from those respondents, and (3) analyzing the data generated to answer the questions of interest. Survey research is a well-established and respected scientific approach to measuring the behavior, attitudes, and beliefs, and much more of populations of individuals.

Contemporary Uses of Surveys

Never before has so much survey research been done and had so much impact on academic scholarship, on government policy-making, and on business decision-making. One likely reason for this interest is that the cohort running government and business now went to college in the 1960s and 1970s and was the first to learn about survey-based social science during their college years, so they have long-appreciated the value of social science data collection and analysis and especially surveys. Another likely reason is the increasing sophistication and practicality of the method itself. First with the advent of telephone interviewing and now with the introduction of Internet-based surveys, turn-around time for survey data collection has decreased considerably as compared to the face-to-face surveys that began the field. So the fast pace of today's decision-making in government and business can be served well by surveys. All this creates a real sense of value placed on the methodology.

As a result, billions of dollars are spent every year conducting surveys around the world,

including surveys conducted or funded by the federal government, surveys done by academics and businesses, and surveys conducted to provide evidence in court cases. The U.S. federal government is one of the largest producers of survey data. For decades, federal agencies have sponsored ongoing surveys to track contemporary life in America in a wide array of domains such as those listed in the top panel of Table 8. The second panel of Table 8 shows other large-scale survey projects that have been funded by the U.S. government over the years.

Survey data form the basis of many important government policy-making decisions. For example, economists at the Federal Reserve and other such agencies pay close attention to the federal unemployment and inflation rates, both of which are calculated using data from national surveys. Many other federal agencies collect survey data because those data are used in on-going decision-making.

Interest in surveys has been growing in the private sector and the academic world as well. For example, whereas survey data appeared in articles in top academic journals in political science, economics, social psychology, and sociology in 1949 at a rate of 3%, 6%, 22%, and 24%, respectively, these figures had increased by 1994 to 42%, 42%, 50%, and 70%, respectively (Presser, 1984; Saris, Gallhofer, van der Veld, & Corten, 2003). This trend no doubt reflects the facts that (1) surveys are now capable of generating much more interesting social scientific data, via implementation of embedded multifactorial experimental designs (e.g., for conjoint analysis) and complex measurement procedures previously thought to be suitable only for specialized surveys or laboratory settings (e.g., for gauging response latencies or complex features of social networks), (2) systematic and meaningful cross-national comparisons are of increasing interest to social scientists, and (3) social scientists used to doing laboratory studies of college students now want to broaden their empirical contributions by generating data on more heterogeneous and representative samples of the population.

Evidence of Survey Accuracy

Decades of research have shown that the reliability and validity of optimally-collected survey data are generally quite high, and that respondents can be relied upon to provide quite accurate descriptions of their past experiences, behaviors, and opinions. Most visibly, surveys conducted just before U.S. presidential elections predict the actual election vote results very closely (see, e.g., Visser et al., 1996; Visser et al., 2000). Even when there is error in such

survey measurements, the error is not huge in percentage point terms. For example, since 1936, the percent of votes won by the winner has correlated with the Gallup Poll's pre-election prediction of that percentage .85, a nearly perfect association.[40] Likewise, since 1948, the American National Election Study surveys' post-election measurements of the proportions of votes won by the winning presidential candidate have correlated with official government vote counts .92, again nearly perfect.

Additional evidence regarding the accuracy of surveys comes from the Monthly Survey of Consumer Attitudes and Behavior, conducted continuously by the University of Michigan's Survey Research Center since 1970. Each month, a representative national sample of American adults has been asked what they expect to happen to the unemployment and inflation rates in the future (as well as many other topics), and their aggregated answers have predicted later changes in actual unemployment and inflation remarkably well (correlations of .80 and .90, respectively, between 1970 and 1995). This is testimony not only to the aggregated wisdom of the American public but also to the ability of scientific surveys to measure that wisdom accurately.

Table 8: Examples of Federal Surveys

Continuously Funded Surveys		
Name of Survey	Agency Sponsor	Years
Survey of Income and Program Participation	Census Bureau	1984-present
Consumer Expenditure Surveys	Census Bureau	1968-present
Survey of Consumer Attitudes and Behavior	National Science Foundation	1953-present
Health and Nutrition Examination Surveys	National Center for Health Statistics	1959-present
National Health Interview Survey	National Science Foundation	1970-present
American National Election Studies	National Science Foundation	1948-present
Panel Study of Income Dynamics	National Science Foundation	1968-present
General Social Survey	National Science Foundation	1972-present
National Longitudinal Survey	Bureau of Labor Statistics	1964-present
Behavioral Risk Factor Surveillance System	Centers for Disease Control and Prevention	1984-present
Monitoring the Future	National Institute of Drug Abuse	1975-present
Continuing Survey of Food Intake by Individuals	Department of Agriculture	1985-present
National Aviation Operations Monitoring System	National Aeronautics and Space Administration	2002-present
National Survey of Drinking and Driving	National Highway Traffic Safety	1991-present

	Administration	
National Survey of Family Growth	National Center for Health Statistics	1973-present
National Survey of Fishing, Hunting, and Wildlife-Associated Recreation	Census Bureau	1991-present
National Survey of Child and Adolescent Well-Being	Department of Health and Human Services	1997-present
Survey of Earned Doctorates	Science Resources Statistics Program, National Science Foundation	1958-present
National Survey on Drug Use and Health	Department of Health and Human Services	1971-present
Youth Risk Behavior Surveillance System	Department of Health and Human Services	1990-present
National Crime Victimization Survey	Bureau of Justice Statistics	1973-present
Schools and Staffing Survey	National Center for Educational Statistics	1987-present
Educational Longitudinal Survey	National Center for Educational Statistics	2002-present
Current Employment Statistics Survey	Bureau of Labor Statistics	1939-present
Other Major Federal Surveys		
Name of Survey	Agency Sponsor	
National Survey of Distracted and Drowsy Driving	National Highway Traffic Safety Administration	
National Survey of Veterans	Department of Veteran Affairs	
National Survey of Children's Health	Health Resources and Services Administration's Maternal and Child Health Bureau	
National Survey of Recent College Graduates	Science Resources Statistics Program, National Science Foundation	
National Survey of Speeding and Other Unsafe Driving Actions	National Highway Traffic Safety Administration, Department of Transportation	

Reasons to Conduct Surveys

Surveys are usually done for one of three reasons: a) to document the prevalence of some characteristic in a population, b) to compare the prevalence of some characteristic across subgroups in a population, and/or c) to document causal processes that produce behaviors, beliefs, or attitudes. Surveys have several major advantages over many other forms of data

collection. Because scientific surveys involve probability sampling, their results can be used to estimate population parameters.

Designs of Surveys

Surveys can take on a variety of designs, which are suitable for addressing different types of research questions. These designs include cross-sectional, repeated cross-sectional, panel, and mixed designs.

Cross-sectional surveys involve the collection of data at a single point in time from a sample drawn systematically from the population. This design is often used to document the prevalence of particular characteristics in a population. For example, a cross-sectional survey might be conducted to assess the frequency with which people have performed a particular behavior or the number of people who hold a particular attitude or belief.

Cross-sectional surveys allow researchers to assess relations between variables and differences between subgroups of respondents and can also be used to test causal hypotheses in a number of ways. For example, statistical techniques such as two-stage least squares regression (James & Singh, 1978) and path analytic techniques (Baron & Kenny, 1986; Kenny, 1979) can be used to test causal relations. Cross-sectional surveys can also be used to identify moderators of relations between variables, thereby also shedding some light on the causal processes at work (e.g., Krosnick, 1988). In addition, a cross-sectional survey can be used to study the impact of an event if the event occurs in the middle of data collection (e.g., Krosnick and Kinder 1990).

Repeated cross-sectional surveys involve collecting data from independent samples drawn from the same population at two or more points in time. Such data can be used to provide evidence about causality, by gauging whether changes in an outcome variable parallel changes in a purported cause of it. If a hypothesized causal relation exists between two variables, between-wave changes in the purported case should be mirrored by between-wave changes in the outcome variable. Repeated cross-sectional surveys can also be used to study the impact of social events that occurred between the surveys (e.g., Weisberg, Haynes, & Krosnick, 1995).

Panel surveys involve collecting data from the same sample of respondents at two or more points in time and can be used to gauge the stability of beliefs, attitudes, or behaviors over time and to identify the determinants of stability (e.g., Krosnick, 1988; Krosnick and Alwin,

1989). Panel surveys can also be used to test causal hypotheses, either by examining whether individual-level changes over time in a purported case correspond to individual-level changes in an outcome variable over the same period in time or by assessing whether changes over time in the outcome variable can be predicted by prior levels of the purported cause. The latter demonstration provides relatively strong evidence consistent with a causal hypothesis, because the changes in the outcome variable could not have caused the prior levels of the purported cause (see, e.g., Blalock, 1985; Kessler & Greenberg, 1981, on the methods; see Rahn, Krosnick & Breuning, 1994, for an illustration of its application). Panel surveys can also permit strong tests of the effect of social events on people.

Panel surveys entail respondent attrition, also referred to as "panel mortality." Some of the people who provide data during the first wave of interviewing will not participate in subsequent waves. Some respondents will be unreachable during the field period of a wave (e.g., due to extended temporary travel away from home or due to moving to a new residence permanently), and a few will refuse to be interviewed at that time, often because of a temporarily busy schedule. Attrition is undesirable, because it reduces a panel's effective sample size. And attrition is more undesirable if a non-random subset of respondents drop out. However, the literature on panel attrition is actually quite reassuring on this point.

A few past studies have documented instances in which people who were and were not reinterviewed differed from one another in some regard. Groves, Singer, and Corning (2000) showed that respondents interviewed initially who had lower levels of community involvement were less likely to be reinterviewed. Another study showed that reinterviewed respondents were less aggressive and more dependent than non-reinterviewed individuals, although the two groups did not differ on many other personality dimensions, such as deference, achievement, nurturance, endurance, dominance, autonomy, and the tendency to offer assistance (Lubin, Levitt & Zuckerman 1962). And Sobol (1959) found that respondents who dropped out of a panel after the first interview were less interested in the topic of the survey than respondents who were successfully reinterviewed, though these groups did not differ in terms of many other variables. But these studies are the exceptions rather than the rule.

Most studies of attrition found little or no sample composition changes as individual panel members dropped out of the active samples. For example, Zagorsky and Rhoton (1999) found very small differences in sample characteristics in the National Longitudinal Survey, even

after half the sample had dropped out. Fitzgerald, Gottschalk, and Moffitt (1998a, 1998b) reached the same conclusion about the impact of 50% drop-out from the Panel Study of Income Dynamics Sample (see also Becketti, Gould, Lillard, & Welch 1988). Falaris and Peters (1998) found no notable effects of attrition on covariation between items in the National Longitudinal Surveys of Labor Market Experience (see also Zabel 1998; Ziliak & Kniesner 1998). And Clinton (2001) likewise found very small differences in attitudinal and demographic variables distributions between people who remained in a long-term panel and those who dropped out. And Traugott (1990) found that as the number of times people were interviewed about politics increased up to four, the respondents who dropped out were slightly less interested in politics and voted slightly less frequently than the respondents who were successfully reinterviewed, but these attrition-related biases were again very small (for similar results, see Price and Zaller 1993; Rahn, Krosnick, & Breuning 1994). Thus, panel attrition seems not to pose a major threat.

A second methodological issue in panel research is the possibility that interviewing people repeatedly may change them and thereby make the sample less representative of the larger population to which investigators wish to generalize. There are a variety of such literatures, and each raises its own considerations. But again, they are reassuring for the most part and helpful because they point to the most likely ways in which conditioning might occur, thereby allowing researchers to look out for, and correct for, conditioning if it occurs.

A number of studies have found either no evidence of panel conditioning effects or very small effects. For example, Cordell and Rahmel (1962) found that participating in Nielsen surveys on media use did not alter later reports of media use. Himmelfarb and Norris (1987) found that being interviewed on a wide range of topics did not alter people's subsequent reports of mental health, physical health, self-esteem, social support, or life events experienced. Sobol (1959) found minimal differences between panel respondents and fresh cross-sectional samples interviewed at the same time in terms of a range of demographic and attitudinal variables. And Clinton (2001) found very small differences between a reinterviewed panel sample and a fresh cross-sectional sample interviewed at the same time. Willson and Putnam (1982) found in a meta-analysis that answering questions caused attitudes toward objects to become slightly more positive and caused people to perform slightly better on cognitive skills tests. But these effects were quite small and unreliable across studies.

Some studies that documented conditioning effects tested the "stimulus hypothesis"

(Clausen 1968), the notion that interviewing people on a particular topic may induce them to become more cognitively engaged in that topic subsequently. For example, Bridge, Reeder, Kanouse, Kinder, Nagy, and Judd (1977) found that interviewing people about health caused a subsequent increase in the importance that these people attached to good health. Likewise, Willson and Putnam (1982) found that interviewing caused slightly increased information gathering about the topic. Kraut and McConahay (1973) reported a field experiment showing that interviewing people about politics prior to an election increased the rate at which these people turned out to vote in a subsequent election two months later. Yalch (1976) found the same result in an experiment in Chicago, and Granberg and Holmberg (1992) found the same effect in an experiment in Sweden. However, Mann (2005) found no effect at all of being interviewed about politics on turnout in a subsequent election, so such an effect is not inevitable. If repeated interviews with panel members touch on a wide variety of topics, each wave may blend in with memories of prior waves via what psychologists call "retroactive interference," thus minimizing the likelihood of stimulated interest in any one topic.

Other studies have documented effects of asking people just one question about their behavioral intentions (see, e.g., Fitzsimons & Morwitz 1996; Morwitz, Johnson, & Schmittlein 1993; Ofir & Simonson 2001; Sherman 1980; Spangenberg & Greenwald 1999). For example, Greenwald et al. (1987) showed that asking some survey respondents whether they would vote in an election in the next few days increased the rate at which these people actually voted. And Gregory, Cialdini, and Carpenter (1982) showed that asking people about being cable TV subscribers increased the rate at which these people actually subscribed to cable TV subsequently. Therefore, it seems that asking respondents about their behavioral intentions may influence the likelihood of inducing these behaviors.[41]

Interestingly, membership in a long-term panel survey may be beneficial to the quality of data collected because of what psychologists call "practice effects." The more a person performs any task, the more facile and effective he or she becomes at doing so. In our case, the tasks of interest include question interpretation, introspection, recollection, information integration, and verbal reporting (see Tourangeau, Rips, & Rasinski 2000). Despite the proliferation of surveys in America, the average American has participated in only one survey during the past 12 months (Krosnick & Bullock, 2005). This means that the cognitive processes entailed in survey responding are relatively rarely performed by most Americans, though of course everyday

conversation entails those very same cognitive tasks, so we are all quite practiced at them outside of the survey context. Nonetheless, Chang and Krosnick (2002) found that respondents who had more prior experience answering survey questionnaires provided survey responses that were more accurate. Therefore, prior interviews may enhance the accuracy of measurement in subsequent waves of a panel study.

Researchers can capitalize on the strengths of each of the designs discussed above by incorporating both cross-sectional and panel surveys into a single investigation. If, for example, a researcher is interested in conducting a two-wave panel survey but is concerned about carry-over effects, he or she could conduct an additional cross-sectional survey at the second wave. That is, the identical questionnaire could be administered to both the panel respondents and to an independent sample drawn from the same population. Significant differences between the data collected from these two samples would suggest that carry-over effects were, in fact, a problem in the panel survey. In effect, the cross-sectional survey respondents can serve as a "control group" against which panel survey respondents can be compared.

Evidence testing causal hypotheses can also be produced with surveys by implementing experiments in the questionnaire. If respondents are randomly assigned to "treatment" and "control" groups that are asked different versions of a question or question sequence, differences between the two groups can then be attributed to the treatment.

Total Survey Error

A high level of accuracy can be achieved if optimal procedures are implemented to conduct a survey, and departures from such procedures can significantly compromise the accuracy of a survey's findings. Necessary features include drawing a representative sample of the population, taking extensive steps to collect data from as many sampled people as possible, optimizing the choice of survey mode to achieve accurate measurements, asking questions that are easily comprehensible and do not entail biased wording or format, weighting results to correct for unequal sampling probabilities, and much more.

Usually, researchers have a fixed budget and must decide how to allocate those funds in order to maximize the quality of their data. This sort of decision-making can be made usefully using a perspective first proposed by Hansen and his colleagues (e.g., Hansen & Madow, 1953). According to this approach, a research can think about survey design issues within an explicit

cost-benefit framework geared toward helping researchers make design decisions that maximize data quality within budget constraints. This approach to survey methodology, known as "total survey error" perspective (cf. Dillman, 1978; Fowler, 1988; Groves, 1989; Lavrakas, 1993).

The total survey error perspective recognizes that the ultimate goal of survey research is to accurately measure particular constructs in a sample of people who represent the population of interest. In any given survey, the overall deviation from the ideal is the cumulative result of several sources of survey error. Specifically, the total survey error perspective disaggregates overall error into four components: coverage error, sampling error, nonresponse error and measurement error. *Coverage error* refers to the bias that can result when the pool of potential survey participants from which a sample is selected does not include some portions of the population of interest. *Sampling error* refers to the random differences that invariably exist between any sample and the population from which it was selected. *Nonresponse error* is the bias that can result when data are not collected from all of the members of a sample. And *measurement error* refers to all distortions in the assessment of the construct of interest, including systematic biases and random variance that can be brought about by respondents' own behavior (e.g., misreporting true attitudes, failing to pay close attention to a question), interviewer behavior (e.g., misrecording responses, providing cues that lead participants to respond in one way or another), and the questionnaire (e.g., ambiguous or confusing question wording, biased question wording or response options).

The total survey error perspective advocates explicitly taking into consideration each of these sources of error and making decisions about the allocation of finite resources with the goal of reducing the sum of each of the four. Many steps can be taken to reduce these four types of error that do not cost real dollars, so these quality-maximizing steps do not impact the bottom line. But other steps to reduce error do cost money, and the more money a researcher spends to reduce one type of error, the less money he or she has available to reduce other types of error. Researchers should make such tradeoffs explicitly, recognizing the opportunity costs they pay when making a particular move to maximize quality in a particular way, selecting approaches likely to yield the biggest bang for the buck spent.

Sampling

When designing a survey's sample, the population (the complete group of elements to

which one wishes to generalize findings obtained from a sample) must be defined, and the subset of elements (the individual unit about which information is sought) in the population to be interviewed must be selected. The sampling frame is the complete list of all of the elements in the population from which the sample is drawn. These decisions decision have important implications for the results of the survey (see, e.g., Laumann, Michael, Gagnon, & Michaels, 1994).

There are two broad classes of sampling methods: nonprobability sampling and probability sampling. *Nonprobability sampling* refers to selection procedures in which elements are not randomly selected from the population or in which some elements have zero or unknown probabilities of selection. Probability sampling refers to selection procedures in which elements are randomly selected from the sampling frame and each element has a known, nonzero chance of being selected. The latter approaches do not require that all elements have an equal probability of being selected, nor do they preclude some elements from having a certain (1.00) probability of selection. However, they do require that the selection of each element must be independent of the selection of every other element. Unlike nonprobability sampling, *probability sampling* allows researchers to be confident that a selected sample is representative of the larger population from which it was drawn. When the probabilities that some elements will be selected are not known or zero, there is no way to know whether the sample is representative of the population. Therefore, generalizing beyond the specific elements included in the sample is only warranted when probability sampling methods are used.

Probability sampling also allows researchers to precisely estimate the amount of variance present in a given data set that is due to sampling error. Sampling error refers to the discrepancy between the sample data and the true population values that are attributable to random differences between the sample and the sampling frame. When probability sampling methods are used, sampling error can be estimated. This estimate reflects the magnitude of uncertainty regarding obtained parameter estimates and is typically expressed in terms of the standard error of an estimate or the variable of sample estimates around the true population value, assuming repeated sampling.

Therefore, the best survey designs (and virtually all scientific surveys) use some form of probability sampling, such as simple random sampling, systematic sampling, stratified sampling,

or cluster sampling. Simple random sampling is the simplest form of probability sampling. With this method, elements are drawn from the population at random, and all elements have the same chance of being selected.

Systematic sampling is a slight variant of simple random sampling that is more convenient to execute (e.g., see Ahlgren, 1983). Systematic sampling requires that all elements in the population be known and listed. Then, based on number of elements in the population and the sample size desired, every nth (where n is most commonly the number of elements in the population divided by the desired sample size) element is selected. This method results in a probability sample only if the order of the elements in the list is random.

Stratified sampling is a variant of random and systematic sampling where the sampling frame is divided into subgroups (i.e., strata), and the sampling process is executed separately on each stratum (e.g., see Ross, 1988; Stapp & Fulcher, 1983). For example, the sample might be separated into strata by geographic region (state or county) or by a respondent characteristic such as race or education.

Cluster sampling is a form of probability sampling that is used primarily in face-to-face surveys when a population is distributed over a broad geographic region. In such an instance, the sample is also broadly geographically dispersed and it is impractical (and very expensive) to contact sample members (perhaps multiple times). As a result, researchers draw a sample with elements in groups (called “clusters”) rather than one-by-one (e.g., see Roberto & Scott, 1986; Tziner, 1987). Then all elements within a cluster are sampled. Cluster sampling can also be done in multiple stages, with two or more sequential steps of random sampling; this is called multistage sampling (e.g., see Himmelfarb & Norris, 1987).

Surveys that use nonprobability sampling methods, such as haphazard sampling, purposive sampling, snowball sampling, and quota sampling do not have the advantages of probability sampling. Haphazard sampling includes methods in which respondents in a survey are selected solely on the basis of convenience (also sometimes called convenience sampling). One example of haphazard sampling would be a survey in which respondents are volunteers obtained via advertising.

Another method of nonprobability sampling is purposive sampling, which involves haphazardly selecting members of a particular subgroup within a population. In these sorts of

studies, the respondents who are selected have a particular characteristics of interest to the researcher. However, the respondents chosen are not *representative* of the subgroup of interest.

Snowball sampling is a variant of purposive sampling where a few members of a subpopulation are located and each is asked to suggest other members of the subpopulation for the researcher to contact. Used most frequently with samples of very rare subpopulations, this method results in a sample of respondents who are not independent from one another and not representative of the subgroup.

The best-known form of nonprobability sampling is quota sampling, which involves selecting members of various subgroups of the population to assemble a sample that accurately reflects known characteristics of the population. Predetermined numbers of people in each of several categories are recruited to accomplish this. For example, proportions of men and women or members of different racial and ethnic groups are selected so as to match the proportions in the population.

Coverage error occurs when the sampling frame excludes some portion of the population. For example, telephone surveys via random digit dialing (RDD) usually exclude households without land line telephones. And both telephone and face-to-face surveys exclude homeless people and those in institutional living situations.

Coverage and sampling error can be minimized by using methods that minimize differences between the sampling frame and the population and by surveying large samples. However, the relation between sampling error and sample size is not linear. A moderate sample size reduces sampling error substantially in comparison with a small sample size, but further increases in sample size produce smaller and smaller decrements in sampling error. Thus, from a total survey error perspective, researchers should recognize that minimizing sampling error by increasing sample size has decreasing marginal returns, so beyond a moderate sample size, the funds necessary to produce a large sample may be better spent reducing other types of error.

Nonresponse Error

Nonresponse error occurs when data are not collected from all of the eligible elements that have been sampled. Nonresponse can occur either because sampled elements are not contacted (e.g., if no one is ever home at a household) or because members of sampled households decline

to participate.

The response rate for a survey is the proportion of eligible members of the selected sample from whom data were in fact collected. In almost all surveys, the response rate is less than 100%. The lower the response rate is, the greater the risk that the sample is not representative of the population. That is, a survey with a response rate of nearly 100% will represent all types of people in proportions closely mirroring the population. As the response rate decreases, it becomes increasingly possible (though not necessarily so) that some types of people are under- or over-represented in the sample, as compared to the population. Such deviations are referred to as "non-response bias" in a sample, that is, bias in the sample due to certain types of people failing to participate.

To maximize response rates and thereby minimize the potential for non-response error, researchers implement various procedures to encourage sampled individuals to participate. For example, the field period during which researchers attempt to contact potential respondents can be lengthened (e.g., Groves & Lyberg 1988; Keeter et al. 2000), the number of times an interviewer can try to contact a household member can be increased (Merkle, Bauman, & Lavrakas, 1993; O'Neil, 1979), financial incentives can be offered to individuals to thank them for their participation (e.g., Singer et al., 1999; Singer, van Hoewyk, & Maher, 2000), letters can be mailed to households in advance informing residents of the legitimacy of the forthcoming survey request (e.g., Camburn et al., 1995; Link & Mokdad 2005), and the questionnaire can be kept as short as possible (e.g., Collins et al. 1988).

All of these strategies for increasing response rates have been found to do so in at least some studies in which these factors were considered one-by-one. However, some strategies, such as sending advance letters, may not always be successful because they not only provide reassuring information to respondents, but they also give advance notice that interviewers will try to contact respondents, and respondents may use this knowledge to avoid being interviewed. Similarly, leaving messages on potential respondents' answering machines does not necessarily increase response rates.

Low response rates increase only the potential for nonresponse error, because nonresponse error is a function of two variables: the response rate and the size of the difference between respondents and nonrespondents. If respondents and nonrespondents do not differ

substantially, response rates will be unrelated to nonresponse bias. That is, a low response rate does not guarantee the presence of non-response bias. It is possible to conduct a survey with a response rate of 20% and end up with data that describe the population quite accurately.

An accumulating number of publications show that as long as a representative sample is scientifically drawn from the population and thorough, professional efforts are made to collect data from all selected potential respondents, variation in response rates from a low of about 20% to a high of about 65% is not associated with an increase in the accuracy of the survey's results (Curtin, Presser, & Singer, 2000; Holbrook, Krosnick, & Pfent, in press; Keeter, Miller, Kohut, Groves, & Presser, 2000). Furthermore, although many surveys manifest substantial non-response error, there is little evidence that the observed amount of nonresponse error is related to the response rate for the survey.

Researchers concerned about non-response error can use one of three approaches for assessing the extent to which their data are distorted due to such error. The first approach involves making aggressive efforts to contact and interview a randomly selected sample of people who refused to participate in the survey. A second strategy divides individuals who were interviewed into two groups: those who were easy to reach and those who were difficult to reach. If we assume that that the latter are more similar to nonrespondents than are the former, the magnitude of non-response error can be gauged by comparing easy to reach and hard to reach respondents. A third strategy involves using information about respondents and nonrespondents (e.g., household address or telephone number) to information about all households in the same communities. This allows researchers to compare some characteristics of respondents and non-respondents, to gauge the magnitude of differences that appear. Smaller differences are more reassuring about the accuracy of the obtained data.

Question Design

Overview. All surveys include questions, and a series of decisions must be made to achieve optimal designs of those questions, maximizing the reliability and validity of the data collected. First, a researcher must decide if each question will be open-ended to closed-ended. For closed-ended questions, a researcher interested in obtaining rank orders of objects must decide whether to ask respondents to report those rank orders directly or to rate each object separately, so the researcher can infer a rank order. If respondents are asked to rate objects, the

researcher must decide how many points to put on the rating scale and whether and how to label the points with words and/or numbers. The researcher must decide the order in which response options will be offered and whether respondents should be explicitly offered the option to say they "don't know" the answer to a question or have no opinion on the matter in question. Once all the questions are written, they must be put into some order for administration. Researchers must also decide how to optimize measurement on sensitive topics, where social desirability response bias may lead respondents to intentionally misreport answers in order to appear more respectable or admirable.

The structure, wording, and order of questions in questionnaires has traditionally been viewed as "an art, not a science," in the words of Princeton University psychologist Hadley Cantril (1951, p. vii) over five decades ago. And in his book *The Art of Asking Questions*, published the same year, Stanley Payne (1951) cautioned that "the reader will be disappointed if he expects to find here a set of definite rules or explicit directions. The art of asking questions is not likely ever to be reduced to some easy formulas" (p. xi). Thirty years later, Sudman and Bradburn (1982) agreed, saying that "no 'codified' rules for question asking exist" (p. 2). Sampling and data analysis are indeed guided by such rules that are backed by elaborate theoretical rationales. But questionnaire design has been thought of as best guided by intuition about how to script a naturally flowing conversation between a researcher and a respondent, even if that conversation is sometimes mediated by an interviewer.

Experienced questionnaire designers have followed some conventions over the years, but those conventions varied enough from individual to individual and from discipline to discipline to suggest that there are few universally accepted principles. If a questioning approach seems to work smoothly when respondents answer a questionnaire, then many researchers presume it will probably yield sufficiently useful data.

In recent years, however, it has become clear that this is an antiquated view that does not reflect the accumulation of knowledge throughout the social sciences about effective question asking. To be sure, intuition is a useful guide for designing questions, and a good questionnaire yields conversations that feel natural and comfortable to respondents. However, intuition can sometimes lead us astray, so it is useful to refine intuition with scientific evaluation. Fortunately, a large body of relevant scientific studies has now accumulated, and when taken together, their

findings clearly suggest formal rules about how best to design questions to maximize the reliability and validity of measurements made by individual questions. However, this work has been scattered across the publication outlets of numerous disciplines, and this literature has not yet been comprehensively and integratively reviewed in a central place. Doing so is the goal of a forthcoming book, *The Handbook of Questionnaire Design* (Krosnick & Fabrigar, forthcoming), and we summarize some of the book's conclusions about the design of rating scales here to provide a flavor of what this literature has to offer.

Rating scale length. When designing a rating scale, one must begin by specifying the number of points on the scale. A large number of studies have compared the reliability and validity of scales of varying lengths (for a review, see Krosnick & Fabrigar, forthcoming). For bipolar scales, which have a neutral or status quo point in the middle (e.g., running from positive to negative), reliability and validity are highest for about 7 points (e.g., Matell & Jacoby, 1971). In contrast, the reliability and validity of unipolar scales, with a zero point at one end (e.g., running from no importance to very high importance), seem to be optimized for somewhat shorter scales, approximately 5 points long (e.g., Wikman & Warneryd, 1990).

Presenting a 7-point bipolar rating scale is easy to do visually but is more challenging when an interviewer must read the seven choices aloud to a respondent, who must hold them in working memory before answering. Fortunately, such 7-point scales can be presented in easier ways without compromising data quality (Krosnick & Berent, 1993). Specifically, such scales can be presented in sequences of two questions that ask first whether the respondent is on one side of the midpoint or the other or at the midpoint (e.g., "Do you like bananas, dislike them, or neither like nor dislike them?"). Then, an appropriately worded follow-up question can ascertain how far from the midpoint the respondents are who settle on one side or the other (e.g., "Do you like bananas a lot or just a little?"). This branching approach takes less time to administer than offering the single seven-point scale all at once, and measurement reliability and validity are higher with the branching approach as well.

Labeling of rating scale points. A number of studies show that data quality is better when all points on a rating scale are labeled with words than when only some are labeled thusly and the others are labeled with numbers or are unlabeled (e.g., Krosnick & Berent, 1993). Furthermore, respondents are more satisfied when more rating scale points are verbally labeled (e.g., Dickinson & Zellinger, 1980). When selecting labels, researchers should strive to select

ones that have meanings that divide up the continuum into approximately equal units (e.g., Klockars & Yamagishi, 1988). For example, "very good, good, and poor" is a combination that should be avoided, because the terms do not divide the continuum equally: the meaning of "good" is much closer to the meaning of "very good" than it is to the meaning of "poor" (Myers & Warner, 1968).

A very common set of rating scale labels used in questionnaires these days was initially developed by Rensis Likert (1932) and assesses the extent of agreement with an assertion: strongly agree, somewhat agree, neither agree nor disagree, somewhat disagree, strongly disagree. Yet a great deal of research shows that this response choice set is problematic because of acquiescence response bias (see, e.g., Couch & Keniston, 1960; Jackson, 1967; Schuman & Presser, 1981). Some people are inclined to agree with any assertion, regardless of its content, and content-free agreement is more common among people with limited cognitive skills, for more difficult items, and for items later in a questionnaire, when respondents are presumably more fatigued (see Krosnick, 1991, 1999). A number of studies now demonstrate how acquiescence can distort the results of substantive investigations (e.g., Jackman, 1973; Winkler, Kanouse, & Ware, 1982), and in a particularly powerful example, acquiescence undermined the scientific value of *The Authoritarian Personality*'s extensive investigation of fascism and anti-Semitism (Adorno, Frankel-Brunswick, Levinson, & Sanford, 1950).

It might seem that the damage done by acquiescence can be minimized by measuring a construct with a large set of items, half of them making assertions opposite to the other half (called "item reversals"). This approach is designed to place acquiescing respondents in the middle of the final measurement dimension but will do so only if the assertions made in the reversals are as extreme as the statements in the original items. To ensure this balance entails extensive pretesting and is therefore cumbersome to implement. Furthermore, it is difficult to write large sets of item reversals without using the word "not" or other such negations, and evaluating assertions that include negations is cognitively burdensome and error laden for respondents, thus increasing both measurement error and respondent fatigue (e.g., Eifermann, 1961; Wason, 1961). Finally, acquiescers presumably end up at the midpoint of the resulting measurement dimension, which is probably not where most belong on substantive grounds. That is, if these individuals were induced not to acquiesce and instead answered the items thoughtfully, their final index scores would presumably be more valid than placing them at the

midpoint.

Most importantly, answering an agree/disagree question always involves answering a comparable rating question in one's mind first. For example, respondents asked to report their extent of agreement or disagreement with the assertion "I am not a friendly person" must first decide how friendly they are (perhaps concluding "very friendly") and then translate that conclusion into the appropriate selection in order to formulate their answer ("disagree" to the original item). It would be simpler and more direct to ask respondents how friendly they are on a rating scale ranging from "extremely friendly" to "not friendly at all." These would be called "construct-specific response options," because they explicitly state levels of the construct being measured (i.e., friendliness). In fact, every agree/disagree rating scale question implicitly requires the respondent to make a mental rating of an object along a continuous dimension representing the construct of interest, so asking about that dimension is simpler, more direct, and less burdensome. Not surprisingly, then, the reliability and validity of rating scales involving construct-specific response options are higher than those of agree/disagree rating scales (e.g., Ebel, 1982; Mirowsky & Ross, 1991; Ruch & DeGraff, 1926; Wesman, 1946). Consequently, it is best to avoid long batteries of questions in the latter format and instead to ask questions using rating scales with construct-specific response options.

Order of rating scale points. People's answers to rating scale questions are sometimes influenced by the order in which the response alternatives are offered. After reading the stem of most rating scale questions, respondents are likely to begin to formulate a judgment with which to answer the question. For example, the question, "How effective do you think the clean-up plan will be?" would induce respondents to generate an assessment of effectiveness before looking at the offered response options. As satisficing respondents read or listen to the answer choices presented, they are likely to settle for the first response option they encounter that is within their "latitude of acceptance." According to Sherif and colleagues (Sherif & Sherif, 1967; Sherif, Sherif, & Nebergall, 1965), people's judgments can be located at single points on continua, but around those points are regions on the continua that people also find acceptable representations of their beliefs. The first such acceptable response option a satisficing respondent hears is the one likely to be selected, thus inducing primacy effects in ratings, which have been observed in many studies (e.g., Belson, 1966; Carp, 1974; Chan, 1991; Matthews, 1929). To prevent this phenomenon from undetectably biasing ratings in a single direction, it is

best to rotate the order of response choices across respondents and to statistically control for that rotation when analyzing the data.

However, this recommendation must be modified in light of conversational conventions about word order. Linguists Cooper and Ross (1975) outlined a series of rules for ordering words in sentences, one of which says that in a list of terms, it is conventional to say the positive or affirmative before the negative (e.g., "plus or minus," "like or dislike," "for or against," "support or oppose"). Similarly, Guilford (1954) asserted that it is most natural and sensible to present evaluative response options on rating scales in order from positive to negative (e.g., "like a great deal" to "dislike a great deal"). Holbrook, Krosnick, Carson, and Mitchell (2000) showed that measurement validity is greater when the order of answer choices conforms to this convention.

Pretesting. Once a questionnaire has been drafted according to the rules above and many other rules for optimizing question design, researchers can conduct pretests to assess whether the questions are understandable and interpreted uniformly and as intended across respondents. Even the most carefully designed questionnaires sometimes include items that respondents find ambiguous or difficult to comprehend. Questionnaires may also include items that respondents understand perfectly well, but interpret differently than the researcher intended. Because of this, questionnaire pretesting is conducted to detect and repair such problems.

A number of methods have been developed for pretesting, including conventional pretesting, behavior coding, and cognitive interviewing. Conventional pretesting involves interviewers conducting a series of interviews as would normally be done and asking the interviewers to report any problems with question interpretation or comprehension they observed (see, e.g., Bischoping, 1989; Nelson, 1985). Behavior coding can also be implemented with recordings of conventional pretests, wherein a researcher notes the occurrence of verbal and non-verbal events that happened during the interaction between the interviewer and the respondent that might indicate problems with a question (e.g., Cannell, Miller, & Oksenberg, 1981). Cognitive interviewing involves administering a questionnaire to individuals who restate each question in their own words and then “think aloud” while answering, verbalizing whatever considerations come to mind as they formulate their responses (e.g., Forsyth & Lessler, 1991). Researchers can also ask specific questions about the questions, to see how respondents interpret them. These pretesting methods all provide insight into the cognitive processes by which

respondents answer questions, which presumably provides insight into the way each item is comprehended and the strategies used to devise answers.

Each of these pretesting methods has advantages and disadvantages. When resources are available, researchers can use multiple pretesting methods to assess newly designed questionnaires. Different methods are likely to identify different types of problems (see Presser et al., 2004).

Mode of Data Collection

Survey data can be collected in one of four primary modes: mail, telephone, face-to-face, and Internet. Telephone and face-to-face surveys are administered by interviewers, whereas mail and Internet surveys involve self-administered questionnaires. Although it might seem that data can be collected equivalently in these different modes, mode choice can in fact produce notable differences in survey findings. So mode choice must be made carefully in light of each project's goals, budget, and schedule.

Face-to-face interviews are often conducted in respondents' homes by a large staff of professional field interviewers. Face-to-face interviews typically involve primarily oral presentation of questions, sometimes supplemented by visual aids. Interviewers used to record respondents' answers on paper questionnaires, but today's face-to-face interviewers often implement computer assisted personal interviewing (CAPI), wherein they read questions displayed on a laptop computer screen and type respondents' answers directly into the laptop. Such computer-assisted technology facilitates the implementation of complex skip patterns (which question a respondent is asked next is determined by his or her answers to prior questions) and implementation of experiments (n which respondents are randomly assigned to be asked different versions of a question).

Telephone interviews are also often conducted when respondents are in their homes. Telephone surveys routinely involve computerized data collection (CATI, which is computer-assisted telephone interviewing), which is easy to implement in a centralized telephone facility. Interviewers conduct telephone surveys by reading questions from a computer screen, and responses are entered directly into the computer. Telephone interviews almost exclusively involve oral presentation of questions and no visual presentation of materials (although some researchers have mailed printed materials to respondents in advance of a telephone interview).

Mail questionnaires and other paper-and-pencil self-administered questionnaires involve visually presenting questions to respondents, having them complete the written questionnaires by writing on them, and return them to the researchers. These questionnaires are most often mailed or dropped off to individuals at their homes, and they may also be hand-delivered in intercept surveys, in which people are stopped in a public place (e.g., when leaving voting booths on election day).

Although self-administered questionnaires have traditionally been provided on paper, computer assisted self-administered interviewing (CASI) has become more common, both as part of a face-to-face survey and as an independent methodology. In this mode, each question appears on a computer screen, and the respondent types in his or her answer, after which the next question appears on the screen. This technology facilitates the use of complex skip patterns and experiments in surveys and allows respondents to answer privately. However, this method also requires that researchers provide respondents with computer access if they do not already have it. Another technology, audio computer-assisted self-administered interviewing (ACASI, involves a form of CASI in which a computer "reads" questions aloud to respondents, who listen on headphones or over the telephone and type their answers on computers or telephone keypads.

In Internet surveys, respondents answer questions online, typically using computers in their homes. Internet surveys can be static, in which the whole questionnaire is shown on a single page, similar to a paper-and-pencil questionnaire. Dynamic or interactive Internet surveys typically present questions one at a time to respondents and thereby facilitate the use of complex skip patterns.

Each survey mode has advantages and disadvantages. When choosing a mode for a particular survey, researchers must consider cost, characteristics of the population, sampling strategy, desired response rate, question format, question content, questionnaire length, length of the data-collection period, availability of facilities, and other considerations. The specific choice of mode should be driven by the purpose of the research as well as the resources available to implement it.

Aspects of the population, including literacy, telephone coverage, familiarity and access to computers, and respondents' motivation, may be important in the decision about mode. Literacy is necessary for self-administered questionnaires, and familiarity with computers is

helpful if computer assisted self-interviewing is to be implemented. Broad telephone coverage of the population is necessary when conducting a survey by telephone. Similarly, Internet access is important if an Internet survey is to be conducted.

When respondents are recruited face-to-face in their homes, this minimizes coverage error, because the resulting sample includes respondents who live in households without working telephones. Coverage error is larger in Random Digit Dial (RDD) telephone surveys, because they exclude respondents without telephones and those with only cell phones. Coverage error for mail surveys depends upon the quality of the list of households that is used as the initial sample frame.

Although probability sampling is possible in all survey modes, mode is also related to the ease with which probability sampling can be implemented and the form it will take. Telephone and face-to-face surveys routinely use probability household sampling strategies, such as random digit dial (RDD) samples and area probability samples. Mail and other self-administered surveys are more commonly used when a list of the entire population is available (e.g., surveys of an organization's membership).

In some Internet surveys, nonprobability sampling methods are used (e.g., inviting individuals to opt in through advertisements on websites), and quotas can be imposed or weighting can be done to improve the apparent representativeness of the sample (e.g., people invited to complete a questionnaire can be selected from a pool of volunteers in numbers that match the population in terms of demographic characteristics). These strategies do not yield scientific samples of the population of interest and therefore run the risk of being unrepresentative of the population (Malhotra & Krosnick, in press).

Some researchers have implemented probability sampling (e.g., RDD telephone calls) to recruit respondents to complete questionnaires weekly via the Internet. Internet access is provided to respondents who do not having it, avoiding one primary source of coverage error in Internet surveys. The possibility of using computers as such was pioneered by Willem Saris, Professor at the University of Amsterdam, who created the Dutch "Telepanel" in the 1970s by recruiting a representative sample of households and giving them computers and modems. This model was adopted by the U.S. firm Knowledge Networks (they have recruited respondents via RDD telephone calls) and by firms in other countries as well (e.g., TNS Forsa in Germany).

Recently, CentERdata, an institute for applied economic research and survey research for the social sciences located at Tilburg University in the Netherlands, set up the CentERpanel, an Internet panel with telephone recruitment in which respondents without Internet access were given technology with which to enter their responses via their televisions.

Mode also influences the response rates achieved in a survey, with face-to-face surveys typically achieving the highest response rates. Telephone surveys achieve somewhat lower response rates, and self-administered mail surveys achieve low response rates unless a sequence of multiple contacts are implemented at considerable cost and with considerable implementation time (see Dillman, 2006).

The choice of mode may also be influenced by the types of information and questions researchers wish to present. If a survey includes open-ended questions, face-to-face or telephone interviewing is preferable because interviewers can probe incomplete or ambiguous answers to ensure the usefulness and comparability of data across respondents. If complex information will be presented as part of the survey, face-to-face interviewing or paper and pencil or Internet questionnaires allow the presentation of both oral and visual information to maximize comprehension.

The need to ask questions about sensitive topics may also influence researchers' decisions about mode. Self-administered questionnaires and computers provide respondents with a greater sense of privacy and therefore elicit more candid responses than interviewer-administered surveys (e.g., Bishop & Fisher, 1995; Cheng, 1988; Wiseman, 1972). Face-to-face interviewing is likely to elicit more honest answers than telephone interviewing, for at least two reasons. Face-to-face interviewers can develop better rapport with respondents than can telephone interviewers. Furthermore, private response methods, such as CASI, can be easily implemented as part of CAPI face-to-face interview by turning the laptop computer around to the respondent and allowing him or her to enter answers about sensitive topics privately.

Face-to-face data collection permits the longest interviews, an hour or more. Telephone interviews are typically quite a bit shorter, usually lasting no more than 30 minutes, because respondents are often uncomfortable staying on the phone for longer. With self-administered questionnaires, response rates typically decline as questionnaire length increases, so they are generally kept even shorter.

Telephone and Internet surveys can be completed in very short field periods, often within a matter of days (though at the cost of lower response rates). In contrast, distributing questionnaires by mail requires significant amounts of time, and follow-up mailings to increase response rates further increase the overall turnaround time. Similarly, face-to-face interview surveys typically require a substantial length of time in the field.

Face-to-face interviews are usually considerably more expensive than telephone interviews, which are usually about as expensive as self-administered questionnaire surveys of comparable size using methods necessary to achieve high response rates. The cost of Internet data collection from a probability sample is about equivalent to that of telephone RDD interviewing.

These many differences between mode contribute to differences in data quality as well. Face-to-face surveys are apparently the winner on most dimensions. Not only do face-to-face surveys have the highest response rates and are the most flexible in terms of the length of survey interviews possible and presentation of complex information. Face-to-face surveys also acquire more accurate reports than do telephone surveys, because respondents are more willing to be honest and work hard to answer survey questions in the former mode than in the latter mode (Holbrook, Green, & Krosnick, 2003).

Internet surveys allow presentation of complex information, and reporting accuracy appears to be higher in Internet surveys than in telephone surveys (Chang and Krosnick, 2001). There are many reasons for this, including: Internet respondents can pace themselves as they go through a questionnaire and complete sections of the surveys at different times; Internet respondents can read and reread questions and answer choices on a screen instead of having to hold them in wording memory while generating answers; members of Internet panels developed long-term relationships of trust with the researchers. Although response rates from Internet surveys based on initial RDD telephone samples are quite low and have similar coverage error to telephone surveys, such difficulties may be reduced by recruiting probability samples of respondents face-to-face in their homes.

General Conclusions About the Use of Survey Methods

Survey research methods are now quite well understood and used widely to produce huge amounts of valuable information. If researchers conform to best practices, the results of surveys

can continue to feed decision-makers and thereby help to optimize practices implemented in government and in the private sector.

Challenges in Using Surveys For Ecosystem Protection Valuation

Introduction. One useful application of the survey method for EPA is the valuation of ecosystems and ecosystem protection and enhancement. A variety of measurement techniques have been developed to assess the monetary value of ecosystems, and these values can be used as input to cost-benefit analyses in the policy-making process.

When survey studies of this type are conducted, many respondents are asked to place a monetary value on an ecosystem with which they are not especially familiar prior to the survey. But this does not mean that respondents lack a value for the ecosystem in question. Throughout the respondents' lives, their experiences have cumulated into a set of beliefs and attitudes stored in long term memory that are the ingredients of their orientations toward objects they will encounter in the future. Just as a child who has had bad experiences with dogs in the past may immediately have a strong, negative reaction to a new dog that he or she meets, so too may adults who have had extensive experience observing and evaluating ecosystems can have a meaningful reaction to a new ecosystem once he or she learns about it. Therefore, an important component of valuation survey design is to describe the ecosystem as fully as possible and in as unbalanced a way as possible.

To do so means that valuation surveys will be different from the vast majority of other surveys that have been conducted in the past. Whereas such surveys have generally sought to measure beliefs, attitudes, past behavior and more simply by asking a sequence of questions, surveys seeking to value ecosystems must devote a considerable amount of time to educating the respondent about the ecosystem in question. This requires respondents to listen to or read relatively long passages of text and perhaps to observe visual presentations of nonverbal information as well, such as charts, maps, drawings, or photographs.

Conveying a large amount of information. In order for the survey's measurements to be valid, it is important that the survey provide all of the information that respondents want in order to make the judgments being asked of them. And it is important that the information be presented in a way that is understandable to all respondents in a sample, no matter what their background is. To assure that this is so, researchers can begin by conducting qualitative and

quantitative research with pretest respondents, who can be told about the ecosystem and asked to describe their reactions to it and to indicate what other information they would like to have in order to make valuation judgments. It is also possible in such pretesting to ascertain how respondents interpret the information they've been given and what misunderstandings might have occurred in need of correction. Through an iterative process of providing information to a group of people, evaluating reactions, revising the information set, presenting it to another group of people, and evaluating their reactions, the presentation can be refined to enhance widespread understanding and sufficiency of the information set.

As with all surveys, pretesting techniques can include (1) think-aloud interviews, during which respondents describe their thoughts as they digest the presented information and answer the questions asked of them, (2) cognitive interviews with structured probes, during which respondents are asked a series of specific questions about the mental processes they went through as they answered each survey question; and (3) behavior coding, in which respondent and interviewer verbal and non-verbal behaviors during questionnaire administration are coded and analyzed to identify potential problems. Also, at the end of the administration of a valuation questionnaire, respondents can answer knowledge quiz questions, ascertaining what information from the presentation was understood and retained by the respondent properly. It is important to bear in mind, though, that retention of the information is not necessary for that information to have had sensible impact on judgments at the time it was presented (Hastie & Park, 1986; Lodge, McGraw, & Stroh, 1989).

The principles of optimal design can be used to construct graphical displays of information that are understandable to a wide range of respondents (e.g., Kosslyn, 1994; Tufte, 2001). Just as "a picture is worth a thousand words," a great deal of information can be presented to respondents in a single visual display that an interviewer can explain to a respondent or that a respondent can read himself or herself. Thus, respondents can gain a sense of the current state of an ecosystem, its state(s) in the past, the history of events that have unfolded to affect its current state, and the possible trajectories of future events that might influence its state.

In order to accomplish the presentation of a sizable set of information (visual or verbal) to respondents, a variety of techniques can be implemented to maximize the chances adequate comprehension at the time of presentation. The information can be presented in the narrative

form of a story, with all the structure typical of stories that engage recipients' attention. For example, by telling respondents that they'll be told about (1) the state of an ecosystem as it used to exist 50 years ago, (2) changes that have occurred to the ecosystem in the intervening years, (3) the causes of those changes, (4) what could be done to reverse those changes, and (5) how this could be implemented, the presentation would present questions to which respondents might want to know the answers, and this desire to resolve puzzles could motivate respondents to attend to and think about the narrative being presented.

Rather than lecturing respondents for a long time period, a questionnaire can maintain respondent engagement by presenting information in relatively small chunks, separated by questions allowing respondents to react briefly to the information they've been given in innocuous ways (e.g., "Had you ever heard of the Golden River before today?"). Respondents can also be asked periodically to verbalize any information that they'd like to have as the story progresses, to allow them to express their cognitive responses to the presentation.

The choice of survey mode has significant implications for ways that the presentation of information about an ecosystem can be designed and optimized. Face-to-face interviewing seems ideal in many regards, because visual displays of any type may be employed, and an interviewer can create a strong sense of interpersonal connection with the respondent as the story is told. Telephone interviewing may permit creating and maintaining the connection between interviewers and respondents, though probably less strongly, and visual displays are usually not possible (unless they have been mailed to respondents in advance of the phone conversation). Computer administration of a questionnaire can include static and dynamic presentation of visual and aural information, and questions can be interspersed, just as would occur during a conversation between a respondent and an interviewer. But it may not be possible to create the strong sense of connection between the respondent and the researcher via computer administration.

Self-administered paper and pencil questionnaires can also implement this same sequence of information presentation. However, it is not possible to present this information to respondents in small chunks, one at a time (because they can look ahead and see the other information to be presented later). If a respondent looks at a large volume of information presented densely on a large set of pages of paper, the sheer volume may be intimidating and/or

dispiriting, thus, minimizing respondent motivation and provoking superficial processing of the information (e.g., skimming instead of reading carefully). Therefore, the self-administered most may be the least desirable for this reason. No matter what mode is employed, it is important to do extensive pretesting of the final instrument to be sure it's working as intended with respondents.

Scale and spatial issues. Because the spatial and temporal scale of ecological systems and services are important dimensions in valuation processes, these dimensions should be reflected in the communication of information about ecological systems and services and in the measurement of value. In order for survey responses to be most useful for developing policies, the scale of the ecosystem or service should be reflected in various aspects of the survey design. For example, the information that respondents receive during the survey interview should, if possible, explicitly describe the scale of a proposed policy or the ecosystem or service for valuation. This is particularly true if the scale is fixed and can be described consistently across presentation of information, evaluation of policies, and valuation of ecological systems and services.

Although some ecological policies may have well-defined scope in terms of physical or temporal scale, in other cases, these may be variables of interest. So instead of specifying these features of a policy, researchers may want to measure whether these features impact respondents' evaluation of the policy. In a survey context, this could be accomplished by manipulating the physical or temporal scales of the proposed policy (in either between-subjects or within-subjects designs). The impact of these features on support for the policy could then be assessed.

Transfer issues. The most effective way to use the survey method to generate valuation evidence applicable to a particular ecosystem protection situation is to design and implement a survey tailored specifically to that situation. However, this requires that time and material resources be devoted to valuation measurement each time EPA must complete a cost-benefit analysis. An alternative, more efficient approach might be to design studies to test whether the findings from a survey about one set of environmental conditions can be extrapolated to a different set of environmental conditions via some sort of mathematical transformation. For example, if a survey measures the value of damage to an ecosystem caused by one oil spill, would it be possible to multiply these damages by three to anticipate the value of comparable

damage caused by three comparable oil spills to three comparable ecosystems? Or is it possible to anticipate the value of the damage caused by a particular oil spill by dividing in half the value of the damage of an oil spill twice as large? Even if such transformations must be done using nonlinear transformations instead of linear ones, it may be possible to conduct parametric research to ascertain the principles according to which such predictions can be made.

Implementing survey research at EPA. The design and conduct of surveys is best done when informed by the large and growing literatures on optimal survey methods. Therefore, it is important that every EPA survey project be implemented at least partly by individuals who are well-versed in these literatures and maintain up-to-date of their most recent developments. This is probably best accomplished by teams of researchers composed partly of full-time EPA employees who specialize in surveys and outside consultants who are experts in survey methods and ecosystem valuation.

OMB approval. OMB clearance is of course required for all EPA surveys, and achieving this clearance requires that a survey meet high standards of quality. In order to maximize the likelihood of approval, it is important that a proposed survey meet a set of criteria: (1) representative sampling of the population of interest with minimal non-coverage error, (2) a very high response rate or a plan to assess the presence of non-response bias, (3) a measuring instrument that has been developed according to optimal design and pretesting practices, and (4) a measurement approach for which a body of empirical evidence documents validity.

Adequate sampling of a population is relatively easy to do for general population samples and becomes increasingly challenging as more specific subpopulations of small size become of interest. Specialized sampling procedures are under development for such situations (e.g., Blair & Blair, 2006; Rocco, 2003). If EPA is interested in conducting surveys of such specialized subpopulations, it may be of value to commission a group of sampling statisticians to recommend a set of procedures to be used in such surveys in the future. Then, these guidelines can be consulted and followed when conducting sampling for such studies in the future.

The literature on response rates is evolving at a very quick pace these days. The literature has two principal strands: (1) exploring the impact of response rates on data accuracy, and (2) exploring the effectiveness of various data collection techniques for enhancing response rates. Although lower response rates are generally not associated with substantially decreased

accuracy, it may be useful for EPA to reanalyze a set of its own past surveys simulating lower response rates and observing the impact that lower response rates would have had on the survey results. If a large set of such simulations can be conducted and yields evidence of little or no change in results, that would be reassuring about potential problems associated with low response rates in future studies (though it would not mean researchers need not strive for high response rates). And if such simulation studies do indicate changes in results associated with reduced response rates, it may be possible to ascertain a typical direction of bias associated with lower response rates. If such systematic bias is detected, it may be possible to build correction algorithms to adjust the results of future surveys to correct for such bias. For example, if lower response rates are typically associated with higher values placed on ecosystems (because people who are pro-environment generally are especially likely to participate in surveys about the environment), it may be possible to develop an approach to adjust observed valuation levels to simulate what would have been obtained if a survey's response rate had been higher.

It might seem obvious that when EPA conducts surveys, all possible steps should be taken to increase the response rate obtained. According to federal convention, that cannot include offering financial incentives to respondents, which have been shown to increase response rates considerably. But EPA can implement a series of other techniques that are thought to enhance response rates, including sending advance letters to potential respondents telling them that they will be contacted for a survey and explaining its legitimacy, leaving messages on answering machines in telephone surveys, lengthening the field period during which data are collected, and many more. However, in order to justify spending money to implement such techniques, it is important to have empirical evidence documenting the effectiveness of these techniques in the context of EPA surveys.

Furthermore, it is important to be sure that efforts to increase the response rate of a survey do not do so by increasing the unrepresentativeness of the sample. For example, telling respondents that a survey is about the environment may increase response rates considerably among people interested in the environment and may decrease response rates by a smaller margin among people who are less interested in the environment, thus enhancing rather than reducing non-response bias. So it may be worthwhile for EPA to conduct studies assessing whether efforts to increase response rates unintentionally decrease sample representativeness.

Another approach to enhancing the likelihood of obtaining OMB approval may be to gather bodies of evidence documenting the effectiveness of particular measurement techniques. For example, there is considerable controversy surrounding the use of contingent valuation (CV) methods in surveys. Yet NOAA's Blue Ribbon Panel concluded that CV is a viable method of valuation. As with any measurement method, CV is likely to produce reliable assessments if implemented optimally but not if implemented suboptimally. It may be of value for EPA to implement evaluation procedures to identify the optimal elements of a CV survey and the optimal way to implement the method, along with evidence directly assessing the validity of measurement obtained from CV surveys. This may help to reassure OBM evaluators of the merit of measurements produced by the method when it is implemented well.

This same sort of developmental work can be conducted with other valuation techniques. For example, conjoint analysis (e.g., Dennis, 1998) has substantial appeal in various circles, but there are also reasons for concern about potential artifacts occurring with this methodology (e.g., see Stevens, Belkner, Dennis, Kittredge, & Willis, 2000; Wainright, 2993). EPA could therefore consider conducting developmental research comparing the validity of conjoint assessments of ecosystem value to the assessments made by other techniques to identify the technique(s) that yield the most valid data.

Finally, new EPA guidelines on surveys suggest that when a survey is expected to obtain a relatively low response rate, investigators should plan to implement techniques to assess sample representativeness. Rather than outlining what such procedures would look like, OMB has left it to investigators to propose and justify such techniques. EPA could therefore commission work to design procedures for this purpose and conduct studies to validate the effectiveness of the procedures.

IRB and ethical considerations. According to federal guidelines for conducting research with humans, surveys that keep respondents' answers confidential (typically by protecting the security of data and stripping identifiers from the data once it has been collected), do not ask sensitive questions, do not involve deception, and do not involve surveying sensitive populations (e.g., children, prisoners, and pregnant women) are excempt from Institutional Review Board (IRB) review. Even surveys that fit these criteria, however, require review by an IRB to ensure that they are exempt. Surveys that require keeping identifiers with data (e.g., for later contact in a

panel study), ask about sensitive topics (e.g., drug use or sexual behavior), involve deception, involve surveying sensitive populations, or those that involve accessing respondents medical records (E.g., to check self-reports of medical testing against official records) may require additional IRB review and procedures to protect respondents.

References

Adorno, T. W., Frenkel-Brunswick, E., Levinson, D. J., & Sanford, R. N. (1950). The authoritarian personality. New York: Harper.

Ahlgren, A. (1983). Sex differences in the correlates of cooperative and competitive school attitudes. Developmental Psychology, 19, 881-888.

Babbie, E. (1990). Survey research methods. Belmont, CA: Wadsworth.

Baron, R. M. & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. Journal of Personality and Social Psychology, 51, 1173-1182.

Becketti, S., Gould, W., Lillard, L., & Welch, F. (1988). The PSID after fourteen years: An evaluation. Journal of Labor Economics, 6, 472-92.

Belson, W. A. (1966). The effects of reversing the presentation order of verbal rating scales. Journal of Advertising Research, 6, 30-37.

Bischoping, K. (1989). An evaluation of interviewer debriefing in survey pretests. In C.F. Cannell, L. Oskenberg, F. J., Fowler, G. Kalton, & K. Bischoping (Eds.), New techniques for pretesting survey questionnaires (pp. 15-29). Ann Arbor, MI: Survey Research Center.

Bishop, G. F. & Fisher, B. S. (1995). “Secret ballots” and self-reports in an exit-poll experiment. Public Opinion Quarterly, 59, 568-588.

Blair, E. & Blair, J. (2006). Dual frame Web/telephone sampling for rare groups. Journal of Official Statistics, 22, 211-220.

Blalock, H. M. (1985). Causal models in panel and experimental designs. New York: Aldine.

Bracht, G. H. & Glass, G. V. (1968). The external validity of experiments. American Education Research Journal, 45, 437-474.

Bridge, R. G., Reeder, L. G., Kanouse, D., Kinder, D. R., Nagy, V. T., & Judd, C. M. (1977). Interviewing changes attitudes – sometimes. Public Opinion Quarterly, 41, 56-64.

Camburn, D., Lavrakas, P. J., Battaglia, M. J., Massey, J. T. & Wright, R. A. (1995). Using advance response letters in random-digit-dialing telephone surveys. Proceedings of the American Statistical Association, Section on Survey Research Methods, pp. 969-974. Washington, D.C.: American Statistical Association.

Cannell, C. F., Miller, P. & Oksenberg, L. (1981). Research on interviewing techniques. In S. Leinhardt (Ed.), Sociological methodology (pp.389-437). San Francisco: Jossey-Bass.

Cantril, H. (1951). Public opinion 1935-1946. Princeton, NJ: Princeton University Press.

Carp, F. M. (1974). Position effects on interview responses. Journal of Gerontology, 29, 581-587.

Chan, J. C. (1991). Response order effects in Likert-type scales. Educational and Psychological Measurement, 51, 531-541.

Chang, L., & Krosnick, J. A. (2001). The accuracy of self-reports: Comparisons of an RDD telephone survey with Internet Surveys by Harris Interactive and Knowledge Networks. Paper presented at the American Association for Public Opinion Research Annual Meeting, Montreal, Canada.

Cheng, S. (1988). Subjective quality of life in the planning and evaluation of programs. Evaluation and Program Planning, 11, 123-134.

Clausen, A. R. (1968). Response validity: Vote report. Public Opinion Quarterly, 32, 588-606.

Clinton, J. D. (2001). Panel bias from attrition and conditioning: A case study of the Knowledge Networks Panel. Stanford.

Collins, M., Sykes, W., Wilson, P., & Blackshaw, N. (1988). Nonresponse: The UK Experience. In R. M. Groves, P. B. Biemer, L. E. Lyberg, J. T. Massey, W. L. Nicholls II, & J. Waksberg (Eds.), Telephone Survey Methodology (pp. 213-232).

Cooper, W. E. & Ross, J. R. (1975). World order. In R.E. Grossman, L. J. San, and T. J. Vance (Eds.), Papers from the parasession on functionalism (pp. 63-111). Chicago: Chicago Linguistic Society.

Cordell, W. N. & Rahmel, H. A. (1962). Are Nielsen ratings affected by non-cooperation conditioning or response error? Journal of Advertising Research, 2, 45-9.

Couch, A. & Keniston, K. (1960). Yeasayers and naysayers: Agreeing response set as a personality variable. Journal of Abnormal Social Psychology, 60, 151-174.

Curtin, R., Presser, S., & Singer, E. (2000). The effects of response rate changes on the index of consumer sentiment. Public Opinion Quarterly, 64, 413-428.

Dennis, D. F. (1998). Analyzing public inputs to multiple objective decisions on national forests using conjoint analysis. Forest Science, 44, 421-429.

Dickinson, T. L. & Zellinger, P. M. (1980). A comparison of the behaviorally anchored rating and mixed standard scale formats. Journal of Applied Psychology, 65, 147-154.

Dillman, D. A. (1978). Mail and telephone surveys: The total design method. New York: Wiley.

Dillman, D. A. (2006). Mail and internet surveys: The tailored design method. New York: Wiley.

Ebel, R. L. (1982). Proposed solutions to two problems of test construction. Journal of Educational Measurement, 19, 267-278.

Eiferman, R. R. (1961). Negation: A linguistic variable. Acta Psychologica, 18, 258-273.

Falaris, E. M. & Peters, H. E. (1998). Survey attrition and schooling choices. The Journal of Human Resources, 33, 531-54.

Fitzgerald, J., Gottschalk, P., & Moffitt, R. (1998a). An analysis of sample attrition in panel data: The Michigan Panel Study of Income Dynamics. NBER Technical Working Papers National Bureau of Economic Research, Inc.

Fitzgerald, J., Gottschalk, P., & Moffitt, R. (1998b). An analysis of sample attrition on the second generation of respondents in the Michigan Panel Study of Income Dynamics. The Journal of Human Resources, 33, 300-344.

Fitzsimons, G. J. & Morwitz, V. G. (1996). The effect of measuring intent on brand-level purchase behavior. The Journal of Consumer Research, 23, 1-11.

Forsyth, B. H. & Lessler, J. T. (1991). Cognitive laboratory methods: A taxonomy. In P. Biemer, R. Groves, L. Lyberg, N. Mathiowetz, & S. Sudman (Eds.), Measurement error in surveys (pp. 393-418). New York: Wiley.

Fowler, F. J. (1988). Survey research methods (2nd ed.). Beverly Hills, CA: Sage.

Frey, J. H. (1989). Survey research by telephone (2nd ed.). Newbury Park, CA: Sage.

Granberg, D. & Holmberg, S. (1992). The Hawthorne Effect in election studies: The impact of survey participation on voting. British Journal of Political Science, 22, 240-247.

Gregory, W. L., Cialdini, R. B., & Carpenter, K. M. (1982). Self-relevant scenarios as mediators of likelihood estimates and compliance: Does imagining make it so? Journal of Personality and Social Psychology, 43, 89-99.

Greenwald, A.G., Carot, C. G., Beach, R., & Young, B. (1987). Increasing voting behavior by asking people if they expect to vote. *Journal of Applied Psychology* 72:315-318.

Groves, R. M. (1989). Survey errors and survey costs. New York: Wiley.

Groves, R. M. & Lyberg, L. E. (1988). An overview of nonresponse issues in telephone surveys. In R. M. Groves, P. B. Biemer, L. E. Lyberg, J. T. Massey, W. L. Nicholls II, & J. Waksberg (Eds.), Telephone Survey Methodology (pp. 191-212).

Groves, R. M., Singer, E., & Corning, A. (2000). Leverage-salience theory of survey participation: Description and an illustration. Public Opinion Quarterly, 64, 288-308.

Hansen, M. H. & Madow, W. G. (1953). Survey methods and theory. New York: Wiley.

Hastie, R. & Park, B. (1986). The relationship between memory and judgement depends on whether the judgment task is memory-based or on-line. Psychological Review, 93, 258-268.

Himmelfarb, S. & Norris, F. H. (1987). An examination of testing effects in a panel study of older persons. Personality and Social Psychology Bulletin, 13, 188-209.

Holbrook, A. L., Green, M. C., & Krosnick, J. A. (2003). Telephone vs. face-to-face interviewing of national probability samples with long questionnaires: Comparisons of respondent satisficing and social desirability response bias. Public Opinion Quarterly, 67, 79-125.

Holbrook, A. L., Krosnick, J. A., Carson, R. T., & Mitchell, R. C. (2000). Violating conversational conventions disrupts cognitive processing of attitude questions. Journal of Experimental Social Psychology, 36, 465-494.

Holbrook, A. L., Krosnick, J. A., & Pfent, A. M. (in press). Response rates in surveys by the news media and government contractor survey research firms. In J. Lepkowski, B. Harris-Kojetin, P. J. Lavrakas, C. Tucker, E. de Leeuw, M. Link, M. Brick, L. Japec, & R. Sangster (Eds.), Telephone survey methodology. New York: Wiley.

Jackman, M. R. (1973). Education and prejudice or education and response-set? American Sociological Review, 38, 327-339.

Jackson, D. N. (1967). Acquiescence response styles: Problems of identification and control. In I. A. Berg (Ed.), Response set in personality measurement. Chicago: Aldine.

Jagodzinski, W., Kuhnel, S. M., & Schmidt, P. (1987). Is there a "Socratic effect" in nonexperimental panel studies? Sociological Methods and Research, 15, 259-302.

James, L. R. & Singh, B. H. (1978). An introduction to the logic, assumptions, and the basic analytic procedures of two-stage least squares. Psychological Bulletin, 85, 1104-1122.

Keeter, S., Miller, C., Kohut, A., Groves, R. M., & Presser, S. (2000). Consequences of reducing nonresponse in a national telephone survey. Public Opinion Quarterly, 64,125-148.

Kenny, D. A. (1979). Correlation and causality. New York: Wiley.

Kessler, R. C. & Greenberg, D. F. (1981). Linear panel analysis: Models of quantitative change. New York: Academic Press.

Klockars, A. J. & Yamagishi, M. (1988). The influence of labels and positions in rating scales. Journal of Educational Measurement, 25, 85-96.

Kosslyn, S. M. (1994). Elements of graphic design. New York Freeman.

Kraut, R. E. & McConahay, J. B. (1973). How being interviewed affects voting: An experiment. Public Opinion Quarterly, 37, 398-406.

Krosnick, J. A. (1988). The role of attitude importance in social evaluation: A study of policy preferences, presidential candidate evaluations, and voting behavior. Journal of Personality and Social Psychology, 55, 196-210.

Krosnick, J. A. (1991). Response strategies for coping with the cognitive demands of attitude measures in surveys. Applied Cognitive Psychology, 5, 213-236.

Krosnick, J. A. & Alwin, D. F. (1989). Aging and susceptibility to attitude change. Journal of Personality and Social Psychology, 57, 416-425.

Krosnick, J. A. & Berent, M. K. (1993). Comparisons of party identification and policy preferences: The impact of survey question format. American Journal of Political Science, 37, 941-964.

Krosnick, J. A., & Bullock, J. (2005). Thought piece on survey participation. Paper presented at the conference entitled "New Approaches to Understanding Participation in Surveys," Belmont Conference Center, Elkridge, Maryland.

Krosnick, J. A. & Fabrigar, L. R. (forthcoming). Designing great questionnaires: Insights from psychology. New York: Oxford University Press.

Krosnick, J. A. & Kinder, D. R. (1990). Altering popular support for the president through priming: The Iran-Contra affair. American Political Science Review, 84, 497-512.

Laumann, E. O., Michael, R. T., Gagnon, J. H., & Michaels, S. (1994). The social organization of sexuality: Sexual practices in the United States. Chicago: University of Chicago

Press.

Lavrakas, P. J. (1993). Telephone survey methods: Sampling, selection, and supervision (2nd ed.). Newbury Park, CA: Sage.

Likert, R. (1932). A technique for the measurement of attitudes. New York: Archives of Psychology.

Link, M. W. & Mokdad, A. (2005). Advance letters as a means of improving respondent cooperation in RDD studies: A multi-state experiment. Public Opinion Quarterly, 69, 572-587.

Lodge, M., McGraw, K. M., & Stroh, P. (1989). An impression-driven model of candidate evaluation. American Political Science Review, 83, 399-420.

Lubin, B. E., Levitt, E. & Zuckerman, M. (1962). Some personality differences between responders and nonresponders to a survey questionnaire. Journal of Consulting Psychology, 26, 192.

Malhotra, N. & Krosnick, J. A. (in press). The effect of survey mode on inferences about political attitudes and behavior: Comparing the 2000 and 2004 ANES to Internet surveys with non-probability samples. Political Analysis.

Mann, Christopher B. 2005. "Unintentional Voter Mobilization: Does Participation in Pre-election Surveys Increase Voter Turnout?" *Annals of the American Academy of Political and Social Science* 601:155-168.

Matell, M. S. & Jacoby, J. (1971). Is there an optimal number of alternatives for Likert scale items? Study I: Reliability and validity. Educational and psychological measurement, 31, 657-74.

Matthews, C. O. (1929). Erroneous first impressions on objective tests. Journal of Educational Psychology, 20, 280-286.

Merkle, D. M., Bauman, S. L., & Lavrakas, P. J. (1993). The impact of callbacks on survey estimates in an annual RDD survey. Proceedings of the American Statistical Association, Section on Survey Research Methods, pp. 1070-1075. Washington, D.C.: American Statistical Association.

Mirowsky, J. & Ross, C. E. (1991). Eliminating defense and agreement bias from measures of the sense of control: A 2 x 2 index. Social Psychology Quarterly, 54, 127-145.

Morwitz, V. G., Johnson, E., & Schmittlein, D. (1993). Does measuring intent change behavior? The Journal of Consumer Research, 20, 46-62.

Nelson, D. (1985). Informal testing as a means of questionnaire development. Journal of Official Statistics, 1, 179-188.

Myers, J. H. & Warner, W. G. (1968). Semantic properties of selected evaluation adjectives. Journal of Marketing Research, 5, 409-412.

Nosanchuk, T. A., & Marchak, M. P. (1969). Pretest sensitization and attitude change. Public Opinion Quarterly, 33, 107-111.

Ofir, C. & Simonson, I. (2001). In search of negative customer feedback: The effect of expecting to evaluate on satisfaction evaluations. Journal of Marketing Research, 38, 170-182.

O'Neil, M. J. (1979). Estimating the nonresponse bias due to refusals in telephone surveys.

Public Opinion Quarterly, 43, 218-232.

Payne, S. (1951). The art of asking questions. Princeton, NJ: Princeton University Press.

Presser, S. (1984). The use of survey data in basic research in the social sciences. In C. F. Turner & E. Martin (Eds.), Surveying subjective phenomenon. New York: Russell Sage Foundation.

Presser, S., Rothgeb, J. M., Couper, M. P., Lessler, J. T., Martin, E., Martin, J, & Singer, E., Eds. (2004). Methods for Testing and Evaluating Survey Questionnaires. New York: Wiley.

Price, V. & Zaller, J. (1993). Who gets the news? Alternative measures of news reception and their implications for research. Public Opinion Quarterly, 57, 133-64.

Rahn, W. M., Krosnick, J. A., & Breuning, M. (1994). Rationalization and derivation processes in survey studies of political candidate evaluation. American Journal of Political Science, 38, 582-600.

Roberto, K. A. & Scott, J. P. (1986). Confronting widowhood: The influence of informatl supports. American Behavioral Scientist, 29, 497-511.; Tziner, 1987; Stapp & Fulcher, 1983

Rocco, E. (2003). Constrained inverse adaptive cluster sampling. Journal of Official Statistics, 19, 45-57.

Ross, M. W. (1988). Prevalence of classes of risk behaviors for HIV infection in an randomly selected Australian population. Journal of Sex Research, 25, 441-450.

Ruch, G. M. & DeGraff, M. H. (1926). Correction for chance and “guess” versus “do not guess” instructions in multiple-response tests. Journal of Educational Psychology, 17, 368-375.

Saris, W. E., Gallhofer, I., van der Veld, W., & Corten, I. (2003). A scientific method for questionnaire design: SQP. Amsterdam: University of Amsterdam.

Schuman, H. & Presser, S. (1981). Questions and answers in attitude surveys. San Diego, CA: Academic Press.

Sherif, M. & Sherif, C. W. (1967). Attitudes as the individual’s own categories: The social judgment-involvement approach. In C. W. Sherif and M. Sherif (Eds.), Attitude, ego-involvement and change (pp. 105-139). New York: Wiley.

Sherif, C. W., Sherif, M., & Nebergall, R. E. (1965). Attitude and attitude change: The social judgment-involvement approach. Philadelphia: W.B. Saunders.

Sherman, S. J. (1980). On the self-erasing nature of errors of prediction. Journal of Personality and Social Psychology, 39, 211-221.

Singer, E., van Hoewyk, J., Gebler, N., Raghunathan, T., & McGonagle, K. (1999). The effect of incentives on response rates in interviewer-mediated surveys. Journal of Official Statistics, 15, 217-230.

Singer, E., van Hoewyk, J. & Maher, M. P. (2000). Experiments with incentives in telephone surveys. Public Opinion Quarterly, 64, 171-188.

Sobol, M. G. (1959). Panel mortality and panel bias. Journal of the American Statistical Association. 54, 52-68.

Spangenberg, E. R. & Greenwald, A. G. (1999). Social influence by requesting self-prophecy. Journal of Consumer Psychology, 8, 61-89.

Stapp, J. & Fulcher, R. (1983). The employment of APA members: 1981. American Psychologist, 38, 1298-1320.

Stevens, T. H., Belkner, R., Dennis, D., Kittredge, D., & Willis, C. Comparison of contingent valuation and conjoint analysis in ecosystem management. Ecological Economics, 32, 63-74.

Sudman, S. & Bradburn, N. M. (1982). Asking questions: A practical guide to questionnaire design. New York: Jossey-Bass.

Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). The Psychology of Survey Response. New York: Cambridge University Press.

Traugott, M. W. (1990). Memo to pilot study committee: Understanding campaign effects on candidate recall and recognition. NES Pilot Study Report, No. nes002284.

Tufte, E. R. (2001). The visual display of quantitative information. Cheshire, CT: Graphics Press.

Tziner, A. (1987). Congruency issues retested using Rineman's achievement climate notion. Journal of Social Behavior and Personality, 2, 63-78.

Visser, P. S., Krosnick, J. A., Marquette, J., & Curtin, M. (1996). Mail surveys for election forecasting? An evaluation of the Columbus Dispatch poll. Public Opinion Quarterly, 60, 181-227.

Visser, P. S., Krosnick, J. A., Marquette, J., & Curtin, M. (2000). Improving election forecasting: Allocation of undecided respondents, identification of likely voters, and response order effects. In P. Lavrakas & M. Traugott (Eds.), Election polls, the news media, and democracy. New York, NY: Chatham House.

Wainright, D. M. (2003). More 'con' than 'joint': Problems with the application of conjoint analysis to participatory healthcare decision making. Critical Public Health, 13, 373-380.

Wason, P. C. (1961). Responses to affirmative and negative binary statements. British Journal of Psychology, 52, 133-142.

Weisberg, H. F., Haynes, A. A., & Krosnick, J. A. (1995). Social group polarization in 1992. In H. F. Weisberg (Ed.), Democracy's feast: Elections in America (pp. 241-249). Chatham, NJ: Chatham House.

Weisberg, H. F., Krosnick, J. A., & Bowen, B. D. (1996). An introduction to survey research, polling, and data analysis. Thousand Oaks, CA: Sage.

Welch, W. W. & Wallberg, H. J. (1972). Pretest and sensitization effects in curriculum evaluation. American Educational Research Journal, 7, 605-614.

Wesman, A. G. (1946). The usefulness of correctly spelled words in a spelling test. Journal of Educational Psychology, 37, 242-246.

Wikman, A. & Warneryd, B. (1990). Measurement errors in survey questions: Explaining response variability. Social Indicators Research, 22, 199-212.

Willson, V. L. & Putnam, R. R. (1982). A meta-analysis of pretest sensitization effects in experimental design. American Educational Research Journal, 19, 249-258.

Winkler, J. D., Kanouse, D. E., & Ware, J. E. (1982). Controlling for acquiescence response set in scale development. Journal of Applied Psychology, 67, 555-561.

Wiseman, F. (1972). Methodological bias in public opinion surveys. Public Opinion Quarterly, 36, 105-108.

Yalch, R. F. (1976). Pre-election interview effects on voter turnout. Public Opinion Quarterly, 40, 331–36.

Zabel, J. E. (1998). An analysis of attrition in the Panel Study of Income Dynamics and the Survey of Income and Program Participation with an application to a model of labor market behavior. The Journal of Human Resources, 33, 479-506.

Zagorsky, J. & Rhoton, P. (1999). Attrition and the National Longitudinal Survey's Women Cohorts. Manuscript, Center for Human Resource Research, Ohio State University.

Ziliak, J. P. & Kniesner, T. J. (1998). The importance of sample attrition in life cycle labor supply estimation. Journal of Human Resources, 33, 507-30.

Addressing the question: "People do not hold stable, well-defined tradeoffs for the component attributes of ecosystem valuation. Methods, surveys or other, will inevitably shape the construction of those values." What are the practical and legal implications of this?

Answering this question requires a careful consideration of a large literature stretching back centuries. We offer below a summary of the most relevant points of that literature.

The History of Concerns About Self-Reports.

For centuries, scholars have attempted to understand the causes of human behavior, and during most of that time, theoretical accounts have posited people's behavior to result from their thinking. When developing their theories of the mind, philosophers such as Locke (1959) and Hume (1739) gained insights through introspection – the process of looking inward to describe consciousness, perception, thought. There is even evidence of introspection in the writings of Socrates as well. But it was Wilhelm Wundt who inaugurated modern psychology using introspection as a primary research method (see, e.g., Schultz, 1981; Leahey, 1991; Hothersall, 1984). Wundt's goal was to understand and explain people's conscious experiences of the world around them. And, Wundt presumed, the only method by which this could be studied was through individuals' descriptions of their experiences of external events.

Wundt's application of introspection was innovative because he utilized it only under conditions of strict experimental control of the stimuli to which his subjects were exposed. Wundt systematically varied stimulus conditions, and he observed consequent changes in the experience reported by subjects. Thus, he was able to gain insights into how the processes of perception unfolded while subjects were in the midst of them.

However, introspection was not easily done, in Wundt's opinion and the opinions of his followers. Rather, subjects had to be trained extensively in order to develop sufficient accuracy and precision of self-insights. Numerous studies were conducted in the laboratories of Wundt, Titchener, and others using this method during the last two decades of the nineteenth century and the first two of the twentieth. But introspection soon experienced a sudden and dramatic fall from grace, the victim of a wide range of criticisms. Significant criticisms asserted that (1) introspection interfered with and

altered the very cognitive processes it was intended to reveal (see, e.g., Kant, 1961, 1968; Comte, 1896), there was little agreement among introspectionists regarding their experiences (e.g., Boring, 1953, p. 174), and the agreement that was observed was attributable to extensive training the individuals received, thus standardizing their biases (see, e.g., Comstock, 1921).

Although these criticisms of the introspective method gradually gained popularity over a period of years, their impact peaked quite forcefully and suddenly in 1913, when psychology experienced a revolution: behaviorism, initiated importantly by John Watson. To be an objective science, Watson charged, psychology must focus its attention on observable events and must discard use of hypothetical mentalistic concepts and terms. Thus, the focus shifted from mental events to behavior, both as the independent variable and as the dependent variable. And introspection quickly fell from its place as psychology's preeminent research method. Reinforcing its fall, Skinner (1953, 1957) later argued forcefully that the verbal descriptions people provide of their own internal processes are learned through metaphor and stimulus generalization, not through any insight into actual cognitive processes taking place within the mind.

But the truth is that in the end, as useful as Watson and Skinner's contributions to social science were, their perspective faded from grace relatively quickly, and was replaced by the notion that people did indeed have attitudes, beliefs, preferences, intentions, and other mental constructs that are stored in long-term memory and govern thinking and action. More than 50 years of serious social science has demonstrated that when people are asked to describe their perceptions of the world, their likes and dislikes, the hopes for the future, and more, these reports seem to result sensibly from past experiences and predict future thinking and action. To be sure, people do not behave as the ideal rational actor that some normative theories might posit as optimal for survival. But people do appear able to render judgments that are systematic and responsive to their knowledge and circumstances.

New Skepticism

Interestingly, recent years have seen a resurgence of skepticism about such judgments. The skeptical school of thought has been driven by a large and growing body of evidence showing that the

attitudes and preferences people express are not perfectly stable over time and can be changed if the judgment they are asked to make is presented in a different way. In short, the skeptics say this:

> "If a person has a true, crystallized attitude, then it shouldn't matter how or when I ask him or her about that attitude – he or she should report it identically each time. But instead, people report different attitudes on different occasion when asked the same question, and people report different attitudes in response to questions that are formulated only very slightly differently."

Thus, say the skeptics, if attitude reports are influenced by minor aspects of questionnaire design (e.g., question wording or question order), there is no underlying attitude in people's minds, and responses to survey questions are merely on-the-spot constructions primarily determined by the context in which they are made.

The evidence for this sort of concern includes demonstrations of response choice order effects, question order effects, question framing effects, interviewer effects, and more. We review some of this evidence briefly to document its flavor.

Response choice order effects. A response choice order effect occurs when the distribution of responses to a closed-ended question is affected by the order in which the response options are offered. Primacy effects occur when response options presented early in a list are chosen more frequently than options presented late in the list. Recency effects occur when response options presented late in a list are chosen more frequently than options presented early in the list.

Response order effects can be tested for by randomly assigning some people to receive a question with the response options in one order and the other people to receive the same question with the response options in the opposite order. In one such study, some respondents were asked this question: "Some people say that we will still have plenty of oil 25 years from now. Others say that at the rate we are using our oil, it will all be used up in about 15 years. Which of these ideas would you guess is most nearly right?" The other respondents were asked the same question with the first two sentences reversed. More people selected the "plenty of oil" option when it was presented second than when it was presented first (Schuman & Presser, 1981).

In another such experiment, some respondents were asked “Should divorce in this country to be easier to obtain, more difficult to obtain, or stay as it is now?” The other respondents were asked “Should divorce in this country to be easier to obtain, stay as it is now, or be more difficult to obtain?” The proportion of respondents who said divorce should be more difficult to obtain was lower when it was presented last (Ayidiya & McClendon, 1990).

Question order effects. A question order effect occurs when answers to one question change depending on whether or not another question has been asked before it. For example, in one experiment, some people were asked “Do you think it should be possible for a pregnant woman to obtain a legal abortion if she is married and does not want any more children?” Other people were asked that question after being asked another one first: “Do you think it should be possible for a pregnant woman to obtain a legal abortion if there is a strong chance of serious defect in the baby?” People were less likely to say the married woman should be allowed to obtain an abortion when the birth defect question was asked first than when it was not (Schuman & Presser, 1981).

In another study, some people were asked: “Do you generally think Al Gore is honest and trustworthy?”, and other people were asked that question after being asked “Do you generally think Bill Clinton is honest and trustworthy?” People were less likely to say Vice President Gore was honest and trustworthy after they had rated his running mate than when they rated Vice President Gore without being asked about President Clinton (Moore, 2002).

Question framing effects. Question framing effects occur when synonymous changes in the wording of a question change answers. For example, Dawes (2000) asked some people: “What are the chances that on your next renewal for your car insurance, you will change from your existing provider?” For other respondents, “change from” was replaced with “renew with.” All answers were given on a scale ranging from 0 to 10, where 0 meant no chance at all, and 10 meant they were certain they would.

More people said there was no chance they would change their car insurance than said they were certain they would renew it. In another study, Rasinski (1989) found that more people said we were spending too little money on “improving the condition of blacks” than said we were spending

too little money on "assistance to blacks." Again, two phrases that seemed to be synonyms elicited different responses.

Interviewer effects. Interviewer effects occur when respondents' answers to survey questions are determined partly who interviews them. For example, interviewers' race or gender may influence survey respondents' answers. In one study, White respondents were more likely to say they supported government programs to help African-Americans when being interviewed by an African-American interviewer than when interviewed by a White interviewer (Cotter, Cohen, & Coulter, 1982). In another study, respondents were more likely to say that women are discriminated against in the workplace when interviewed by a woman than when interviewed by a man (Kane & Macaulay 1993).

The Truth About the Evidence: Small, Rare Effects

These and many other such studies have been cited as illustrations of the fact that people have no real opinions, and their answers to questions are shaped by the question wordings or formats or the interviewers who administer them. But this criticism exaggerates and misinterprets the nature of the available evidence in multiple ways. First, it is not the case that every single person in a sample of research participants can be induced to completely reverse his or her statements by subtle manipulations of questions. Instead, the vast majority of the studies cited in this regard simply show statistically significant but relatively small changes in the distributions of responses due to changes in question presentation. In every study, the results are compatible with this portrayal: the vast majority of respondents expressing opinions were unaffected by the manipulations, and only a very small subset of people were influenced. It is also possible that larger groups of people were influenced, but that each of these people was influenced only very slightly. But regardless of the underlying process, the net change in attitude reports has always been small.

Consider, for example, the studies we described above. Schuman and Presser's (1981) experiment about oil supplies found that changing the order of response choices led to a 9 percentage point shift in selection of the "plenty of oil" response in one survey and a 14 percentage point shift in another. Ayidiya and McClendon (1990) found that changing response choice order changed the percentage of people saying divorce should be more

difficult by 11 percentage points. The proportion of people who supported a woman's right to get an abortion "if she did not want any more children" shifted by 13 percentage points depending on the presence of absence of the prior question about a birth defect. The proportion of people saying that Vice President Gore was honest and trustworthy dropped 8 percentage points as the result of asking first about President Clinton in Moore's (2002) study. Whereas 51% of people said there was no chance they would change their car insurance, 41% said they were certain they would renew it (Dawes, 2000). Whereas 36% of respondents said we were spending too little money on "improving the condition of blacks," 28% said we were spending too little money on "assistance to blacks" (Rasinski, 1989). When asked to report whether Blacks are discriminated against, 40% of White respondents said "yes" when being interviewed by a White interviewer, and 52% said "yes" when being interviewed by a Black interviewer (Cotter, Cohen, & Coulter, 1982). And when asked whether women are discriminated against in the workplace, 71% of respondents said so to a male interviewer, and 78% said so to a female interviewer (Kane & Macaulay, 1993). Thus, in studies like these, the effects of questioning manipulations are typically quite small.

A second way in which the skeptics exaggerate their portrayal has to do with the prevalence of questioning effects. It is not the case that these effects happen with all questions or even with most questions. For example, Holbrook et al. (2006) analyzed 548 experiments varying the order of response choices in telephone surveys done by the Gallup Organization. Only 21% of these experiments showed statistically significant response choice order effects, and the average effect of response choice order on the proportion of respondents expressing an opinion was only 2 percentage points. Similarly, Schuman and Presser (1981) found statistically significant response choice order effects in 29% of the experiments they reported. Thus, the vast majority of the time, researchers would have obtained the same or nearly the same results regardless of which order the answer choices were offered in.

The same is true for question order effects as well. Schuman and Presser (1981) examined 113 attitude questions and found that changes in question order produced statistically significant changes in answers in only 7% of them. Likewise, Smith (1991) examined the effect of question order on answers to 944 questions and found statistically significant impact on fewer than 5%. Thus,

researchers would obtain the same results in the vast majority of cases no matter which order the questions were asked in.

In sum, the effects of most manipulations of questions were quite small and do not suggest that a subtle change in the way a question is asked can shift an 80%-20% distribution of opinions to become a 20%-80% distribution. Most often, a synonymous question change does not change answers at all.

The skeptics' claim that attitude reports in surveys are tremendously unstable over time has turned out to be incorrect as well. In a landmark paper published in 1964, Philip Converse (1964) argued that most answers to attitude questions in surveys are really better thought of as "non-attitudes": responses that are completely random choices from among the offered answer choices, pure guesses. The main evidence Converse offered in support of this shocking claim was that people answered a question one way in an interview on one occasion and were very likely to answer differently two years later. However, many researchers have since shown that instability in attitude reports over time is mainly due to two other causes instead: (1) real change in real attitudes over time, in response to new information learned about the issue, and (2) measurement error caused by ambiguous questions and ambiguous answer choice wordings (e.g., Achen, 1975; Krosnick, 1988; 1991).

Another reason not to be discouraged by the critics' evidence is that in many instances, what researchers have claimed were subtle changes in question wording were not in fact subtle at all – they changed the object being evaluated or the nature of the judgment being requested, and respondents' judgments are refined enough, and respondents were attentive enough, so their answers were appropriately responsive to these changes in questions. For example, Newport (2004) described a comparison of two questions asking Americans about the death penalty. One question asked: "Are you in favor of the death penalty for a person convicted of murder?" A second question asked: "If you could choose between the following two approaches, which do you think is the better penalty for murder: the death penalty or life imprisonment with absolutely no possibility of parole?" Whereas 74% of respondents reported that they favored the death penalty in response to the first question, only

53% reported that they preferred the death penalty in response to the second. But to view the difference between these results as evidence that respondents do not have real attitudes is a misinterpretation. The two questions ask for two different judgments: some respondents could answer the first question affirmatively (to indicate that they think the death penalty should sometimes be imposed, though only for a minority of murder convictions) and could select "life imprisonment" in response to the second question (to indicate that they favor life imprisonment for the majority of murder convictions).

Another such example involves surveys asking respondents to evaluate President Clinton after his 1998 speech denying a sexual relationship with Monica Lewinsky. Respondents in several national surveys evaluated President Clinton more negatively when they were asked to evaluate him "as a person" than when they were asked to evaluate his job performance. This reflected meaningful complexity of public attitudes toward President Clinton: people thought more highly of his effectiveness as the nation's leader than they did of his personal conduct. Again, the discrepancy between answers to these questions is not evidence of non-attitudes.

When considering studies that seem to suggest people's opinions are uncrystallized and easily manipulated, another reason to be skeptical involves the participant population involved and the settings in which the data are collected. A large number of studies claiming to show large effects of question wording or formulation have collected data from arbitrary sets of college students who were enrolled in introductory social science courses and were required to participate in the experiments in order to fulfill the course requirements. There are many theory-grounded reasons to suspect that college students may not have preferences as crystallized and informed as older adults who have had decades of experience thinking about collective problems and society (e.g., Sears, 1986). And students fulfilling course requirements have no incentive to carefully consider the implications of the answers they provide on questionnaires, which they know will have no significant influence on anything. Likewise, studies that ask people on their way into a science museum to answer questions based on a few sentences of information without careful consideration or any information about the effects their answers will have may well overestimate the extent of manipulability of those judgments

(Green, Kahneman, & Kunreuther, 1994). Lots of research shows that reasoning is more thoughtful and precise and less manipulable when people believe that their judgments will be consequential (e.g., Lerner & Tetlock, 1999), so such studies should be generalized only very cautiously.

Finally, it is important to recognize that lots of evidence shows that the beliefs and attitudes people express powerfully predict their behavior (see Schuman and Johnson, 1976 for a review). For example, Kim and Hunter (1993) conducted a meta-analysis of 138 estimates of the attitude-behavior relation. The overall correlation between attitudes and behavior was extremely strong (r = .79). Similarly, Krause (1990) meta-analyzed 83 studies and found the correlation to be even slightly stronger (r=.83).

Other evidence regarding the predictive power of attitudes and other survey self-reports can be found in the literature exploring the predictors of vote choice. For example, Schuman and Johnson (1976) reported that between 1936 and 1972, Gallup's pre-election predictions of Presidential election winners' vote chares deviated only 2.4% on average from their votes on election day. Correlations of predictions from attitude surveys and election outcomes provide similar evidence. Since 1936, the percent of votes won by the winner has correlated with the Gallup Poll's pre-election prediction of that percentage .85, a nearly perfect association. Likewise, since 1948, the American National Election Study surveys' post-election measurements of the proportions of votes won by the winning presidential candidate have correlated with official government vote counts .92, again nearly perfect.

Although there is substantial evidence that attitudes powerfully predict behavior, some research has found that attitudes and behavior were not strongly related. But these instances are mostly due to mistakes made by the researchers. One such error has been to assess respondents' attitudes and behavior at different levels of specificity. For example, Fishbein (1966) noted that attitudes are typically measured more broadly than behaviors and argued that a stronger attitude-behavior relation would be observed when attitudes and behavior are measured at the same level of specificity. Many studies have since confirmed the validity of this assertion (e.g., Heberlain & Black, 1976; Katz, 1982; Liska, 1974; Perry, 1976; Weigel & Newman, 1976; Weigel, Vernon, & Tognacci, 1974).

For example, Liska (1974) measured how frequently students had engaged in 8 unethical examination-taking behaviors, including “got help on a take home examination from another student,” “secured a copy of an old examination,” “copied from another person’s paper on an in-class examination,” and others. The frequency with which each respondent had performed each behavior was predicted much more strongly by attitudes toward that behavior than by attitudes toward college cheating in general or toward honesty in interpersonal relations generally (Liska, 1974).

Another mistake made by researchers has been to measure attitudes with only one question instead of using many (e.g., Fishbein & Ajzen, 1974). Answers to any single question measuring an attitude will contain some measurement error due to ambiguity in the meanings of the words and some ambiguity in what people “see” when they look into their hearts and minds to find the answer a question. Although the construct of interest (i.e., the attitude or behavior being measured) influences the answers, answers are also influenced by other factors. No single question perfectly reflects the underlying construct it is designed to measure, so using multiple questions allows researchers to measure opinions more precisely, as shown in many studies documenting much stronger relations between constructs when correcting for measurement error by using multiple questions (Goren, 2004; Hurwitz & Peffley, 1987; Jackson, 1983; Judd & Milburn, 1980; Layman & Carsey, 2002; Moskowitz & Jenkins, 2004; Norpoth & Lodge, 1985)

The lesson from this evidence about measurement error is that it is a mistake to equate a person’s statement of his or her attitude or belief with the true underlying attitude or belief stored in his or her long-term memory. The fact that people’s reports are imperfect because of ambiguities in their internal psychological cues and ambiguity in the words in questions asked to measure those cues does not mean that the underlying construct does not exist, nor does it mean that the underlying construct is so uncrystallized as to be meaningless.

Consider an example using a physical construct such as temperature. Imagine that one wants to measure the temperature of water in a bucket. The thermometers available to conduct this measurement are three years old and heavily-used. When placed in the bucket simultaneously, these three thermometers produce three slightly different readings: 50 degrees, 48 degrees, and 52 degrees. However, this does not mean that we should conclude that the bucket of water has no true

temperature. It means that the thermometers being used to measure the bucket's temperature are imperfect, so their readings contain error. This error is the fault of the measuring tool, not the fault of the water being measured. To conclude that the water has no temperature in this example is equivalent to asserting that people have no preferences, simply because multiple measurements of those preferences with multiple measuring instruments yield different results.

General Conclusion

For these and other reasons, it would be inappropriate to presume that the huge mountain of evidence showing that people express different opinions when asked different questions or at different times means that people do not have any real opinions at all. Nonetheless, it is impossible to generate direct and definitive empirical evidence of the existence of real preferences, because no probe has yet been developed to insert in the brain and confirm the existence of a preference. Therefore, social scientists must settle for indirect and imperfect evidence – we must ask ourselves: "If people do in fact have real preferences that are imperfectly measured by our questionnaires, what empirical patterns would we expect to see?" And if we do in fact observe those empirical regularities, then we can at least sustain our openness to the possibility that people in fact have genuine preferences.

It is important to bear in mind that in this evaluation process, we must not expect the preferences people express will conform to the demands of the ideal homo economicus. We now know that people do not behave optimally according to normative principles of rationality. So there is no reason why their thinking or preferences should obey such laws either. In other words, the plausibility of real preferences existing in people's minds must not hinge on the degree to which those preferences are formed or have effects according to economic theory. Economy theory's plausibility must be tested and refined using objective empirical measurements of people's thinking and action.

How To Make Policy With Imperfect Survey Measurements

As reassuring as the above logic may seem about the plausible existence of real preferences, it does put policy-makers in a potentially awkward position. We have said that the attitudes, beliefs, and preferences that people express are subject to manipulation due to the measuring tool used to assess them. Yet when policy-makers seek to be responsive to public opinion, they will often be

confronted with divergent measurements of public opinion, suggesting that the public thinks different things. So how should a policy-maker select among these various measurements in order to discern "true" public preferences?

The secret to solving this puzzle lies in diagnosing the reasons for the divergence of the measurements. Some assessments that at first seem to tap the same attitude may differ from one another because they are in fact measuring different attitudes. Respondents are sometimes very sensitive to small differences in language, so they may infer meaning in small changes in question wording that lead them to give different responses. Thus, researchers writing survey questions must be very careful about the language that they use, so that they communicate precisely what they want. And policy-makers should be careful when interpreting survey results by paying close attention to the question wordings employed. It would be useful to consider the results obtained by various similar question wordings to bracket the range in which the opinions of interest fall.

In other cases, two measures that seem to tap the same opinion may differ from one another simply because of random measurement error. In such cases, aggregating (e.g., averaging together) multiple measures can help reduce measurement error. Such aggregation cancels out the effects of random measurement error present in any one assessment.

Finally, as we have seen, measurement artifacts (e.g., due to question or response choice order) can produce small but systematic shifts in the distributions of answers. As a result, policy-makers should never presume that a single measurement of public opinion is perfectly precise. For this reason as well, policy-makers should consider a confidence interval around every measurement to reflect the uncertainty inherent in it. For example, if a survey measure suggests that 70% of Americans support a particular policy, it is best to assume that the true proportion of the population who support the policy is somewhere near 70%. It is unlikely to be 95%, but it could well be 65% or 75%. Such confidence intervals can be estimated routinely in surveys, and most often, policy-making does not require greater precision than surveys can afford.

More generally, when and why should EPA survey the public about their values, desires, preferences?

Addressing the questions: "The public likely hold values that arguably should not be part of a normative, regulatory decision (what should be done) -- where, in EPA policy processes, might surveys be used? Advantages/disadvantages at each level

For centuries, democratic theorists have wrestled with the question of how a polity's desires should and do influence policy-making in a democracy. Many have argued that public participation in the democratic process requires an informed and attentive public, but much evidence suggests that the public is not at all well-informed about political matters (see e.g., Delli Carpini & Keeter, 1996). If the public is in fact not well-informed, perhaps they cannot offer thoughtful opinions about such policy matters, so allowing public influence on government decision-making would be irresponsible (de Tocqueville, 1835; Schumpeter, 1950).

However, others have argued that public involvement is a fundamental requirement for a democracy and that public involvement can aid a democratic government in a number of ways (e.g., Berelson, 1952). A democratic government that is not responsive to public will may be doomed to failure, because a government can survive only as long as its citizens believe that they are being adequately represented (e.g., McCann, 1962). Policy-making that is repeatedly contrary to public opinion may therefore eventually produce widespread public dissatisfaction and vocal protest, and eventually a change in the government, either through election of new public officials or perhaps ultimately the collapse of the system of government itself.

Other scholars have also begun to question evidence the extent to which evidence of low levels of political knowledge repudiates the competence of ordinary citizens to participate in the democratic process. It is not clear precisely what knowledge is necessary for people to be effective democratic citizens, or that the questions posed to survey respondents measure that knowledge (Krosnick, 1998; Kuklinski & Quirk, 2002, Lupia, in press). In fact, only recently have scholars begun to articulate the conceptual foundations of citizen competence, explicitly delineating the specific tasks that confront citizens, the criteria by which task performance should be evaluated, the observable indicators of the criteria, and the standards against which the indicators should be evaluated (Kuklinski & Quirk, 2001). Without such systematic analysis, researches have not been

properly equipped to reach any conclusions about whether answers to arbitrary sets of knowledge quiz questions in surveys condemn citizens' ability to inform policy-making.

Some critics have suggested that the volume of political information retained in memory (and therefore available for retrieval during an interview) may vastly underestimate the amount and diversity of information upon which people's political opinions are based (Lodge, Steenbergen, & Brau, 1995). Rather than meticulously storing every piece of political information to which they are exposed, people may simply adjust their attitudes on-line, modifying their views in light of new information, and discard the specific pieces of information that caused the update. That is, having incorporated the information into their relevant opinions, people may choose not to expend the additional effort to retain the information in memory. It may be misleading, therefore, to use tests of political knowledge to draw inferences about the degree to which ordinary citizens hold informed opinions.

Finally, some scholars have begun to challenge the way that political knowledge has been assessed in surveys, suggesting that the "pop quiz" format typically used is misleading regarding the process by which citizens wield political information in consequential judgments and decisions (Prior & Lupia, 2006). Such critics suggest that a citizen's competence is indicated not by how many discrete bits of information are stored in his or her memory and available for instantaneous retrieval, but instead is instantiated by his or her ability and motivation to find and use relevant pieces of information when the task at hands calls for it. Thus, the fact that most people do not have an encyclopedic set of political facts at their fingertips does not mean that the political judgments they make are groundless. When faced with a consequential judgment or decision, people may well seek out and make use of relevant information. Indeed, when provided with both opportunity and motivation, ordinary citizens prove to be quite capable of utilizing political information (Prior & Lupia, 2006). So policy-makers might prefer to measure public preferences in ways (e.g., via Internet surveys) that permit people to pause and seek out any information they need before answering a question.

If one endorses the view that the public should have a voice in government, either because it is necessary for a successful government or because one believes the public is sufficiently informed to do so, the EPA has reason to measure and consider the public's preferences when making policy. This may raise concerns about the amount of information the public has about the issues dealt with by the EPA and the level of understanding brought to issues that have complex policy and scientific implications. In order to address these concerns, one or both of two steps can be taken in measuring public opinion. First, the EPA can conduct surveys that educate survey respondents about a problem or policy before asking for their preferences. Alternatively, in addition to preferences, the EPA can routinely measure the degree of knowledge and thought that people bring to a survey. This latter approach could be used to determine whether the preferences of members of the public who are knowledgeable and thoughtful are substantially different from those who are not. If this is the case, the EPA could conduct statistical simulations to predict what preferences people would express if everyone were fully informed and deeply thoughtful.

TO THE SAB FROM ALLYSON AND JON: We would be happy to offer opinions on when in the EPA policy-making sequence surveys should be done, but to do so, we'd need to know more about the EPA policy-making process than we do now. So some tutoring would be in order.

References

Achen, C. H. (1975). Mass political attitudes and the survey response. American Political Science Review, 69, 1218-1231.

Ayidiya, S. A. & McClendon, M. J. (1990). Response effects in mail surveys. Public Opinion Quarterly, 54, 229-247.

Berelson, B. (1952). Democratic theory and public opinion. Public Opinion Quarterly, 16, 313-330.

Campbell, A., Converse, P. E., Miller, W. E., & Stokes, D. E. (1980). The American Voter. New York: John Wiley and Sons.

Converse, P. E. (1964). The nature of beliefs systems in mass publics. Ann Arbor, MI: The University of Michigan Survey Research Center.

Cotter, P. R., Cohen, J. & Coulter, P. B. (1982). Race-of-interviewer effects in telephone interviews. Public Opinion Quarterly, 46, 278-284.

Dawes, J. (2000). The impact of question wording reversal on probabilistic estimates of defection/loyalty for a subscrition product. Marketing Bulletin, 11, Research Note 1, 1-9.

Delli Carpini, M., & Keeter, S. (1996). What Americans know about politics and why it matters. New Haven: Yale University Press.

de Tocqueville, Alexis. (1835). Democracy in America.

Fazio, R. H. (1986). How do attitudes guide behavior? In Sorrentino, R. M. and Giggins, E. G. (Eds.), Handbook of motivation and cognition (pp. 204-243). New York: Guildford Press.

Fishbein, M. (1966). The relationships between beliefs, attitudes, and behavior. In S. Feldman (ed.), Cognitive Consistency, pp. 202-222. New York: Academic Press.

Fishbein, M. & Ajzen, I. (1974). Attitudes toward objects as predictors of single and multiple behavior criteria. Psychological Review, 81, 59-74.

Goren, P. (2004). Political sophistication and policy reasoning: A reconsideration. American Journal of Political Science, 48, 462-478.

Green, D. P., Kahneman, D., & Kunreuther, H. (1994). How the scope and method of public funding affect willingness to pay for public goods. Public Opinion Quarterly, 58, 49-67.

Heberlain, T. A. & Black, J. S. (1976). Attitudinal specificity and the prediction of behavior in a field setting. Journal of Personality and Social Psychology, 33, 474-479.

Hobbes, T. (1651). Leviathan: The Matter, Form, and Power of a Commonwealth Ecclesiastical and Civil.

Holbrook, A. L., Krosnick, J. A., Moore, D. W., & Tourangeau, R. (2006). Response order effects in categorical questions presented orally: The impact of question and respondent attributes. Unpublished manuscript.

Hurwitz, J. & Peffley, M. (1987). How are foreign policy attitudes structured? A hierarchical model. American Political Science Review, 81, 1088-1120.

Jackson, J. E. (1983). The systematic beliefs of the mass public: Estimating policy preferences with survey data. Journal of Politics, 45, 840-865.

Judd, C. M. & Milburn, M. A. (1980). The structure of attitude systems in the general public: Comparisons of structural equation models. American Sociological Review, 45, 627-643.

Kahneman, D. & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. Econometrica, 47, 263-291.

Kane, E. K. & Macaulay, L. J. (1993). Interviewer gender and gender attitudes. Public Opinion Quarterly, 57, 1-28.

Katz, J. (1982). The impact of time proximity and level of generality on attitude-behavior consistency. Journal of Applied Social Psychology, 12, 151-168.

Kim, M. S. & Hunter, J. E. (1993). Attitude-behavior relations: A meta-analysis of attitudinal relevance and topic. Journal of Communication, 43, 101-142.

Krause, S. J. (1990). Attitudes and the prediction of behavior: A meta-analysis. Paper presented at the Annual Convention of the American Psychological Association. Boston, MA.

Krosnick, J. A. (1988). Attitude importance and attitude change. Journal of Experimental Social Psychology, 24, 240-255.

Krosnick, J.A. (1998). Review of What Americans know about politics and why it matters by M. X. Delli Carpini and S. Keeter. Annals of the American Academy of Political and Social Science, 559, 189-191.

Krosnick, J. A. (1991). The stability of political preferences: Comparisons of symbolic and non-symbolic attitudes. American Journal of Political Science, 35, 213-236.

Kuklinski, J. H., & Quirk, P. J. (2001). Conceptual foundations of citizen competence. Political Behavior, 23, 285-311.

Layman, G. C. & Carsey, T. M. (2002). Party polarization and party structuring of policy attitudes: A comparison of three NES panel studies. Political Behavior, 24, 199-236.

Lerner, J., & Tetlock, P.E. (1999). Accounting for the effects of accountability. Psychological Bulletin, 125, 255-275.

Liska, A. E. (1974). Attitude-behavior consistency as a function of generality equivalence between attitude and behavior objects. Journal of Psychology, 86, 217-33.

Locke, J. (1689/1690). Two treatises of government.

Lupia, A. (in press). How elitism undermines the study of voter competence. In press at Critical Review.

McCann, J. A. (1962). An electorate adrift?: Public opinion and the quality of democracy in Mexico. Latin American Research Review, 38, 60-81.

Moore, D. W. (2002). Measuring new types of question-order effects: Additive and subtractive. Public Opinion Quarterly, 66, 80-91.

Moskowitz, A. N. & Jenkins, J. C. (2004). Structuring political opinions: Attitude consistency and democratic competence among the U.S. mass publics. Sociological Quarterly, 45, 395-419.

Newport, Frank. (2004). Polling matters: Why leaders must listen to the wisdom of the people. New York: Warner Books.

Norpoth, H. & Lodge, M. (1985). The difference between attitudes and nonattitudes in the mass public: Just measurement? American Journal of Political Science, 29, 291-307.

Perry, R. (1976). Attitude scales as behavior estimative devices: Scale specificity and predictive accuracy. Journal of Social Psychology, 100, 137-142.

Petty, R. E. & Krosnick, J. A. (Eds.). (1995). Attitude strength: Antecedents and consequences. Hillsdale, NJ: Lawrence Erlbaum Associates.

Prior, M., & Lupia, A. (2005). What citizens know depends on how you ask them: Experiments on political knowledge under respondent-friendly conditions. Paper presented at the 101st Annual Meeting of the American Political Science Association, Washington, D.C.

Rasinski, K. A. (1989). The effect of question wording on public support for government spending. Public Opinion Quarterly, 53, 388-394.

Schuman, H. & Johnson, M. P. (1976) Attitudes and behavior. Annual Review of Sociology, 2, 161-207.

Schumpeter, J. A. (1950). Capitalism, socialism, and democracy, 3rd edition. New York: Harper and Row.

Sears, D. O. (1986). College sophomores in the laboratory: Influences of a narrow database on social psychology's view of human nature. Journal of Personality and Social Psychology, 51, 515-530.

Smith, T. W. (1991). Context effects in the General Social Survey. In P. B. Biemer, et al. (eds.), Measurement Errors in Surveys. New York: Wiley.

Tversky, A. & Kahneman, D. (1991). Loss aversion in riskless choice: A reference-dependent model. Quarterly Journal of Economics, 106, 1039-1061.

Weigel, R. H. & Newman, L. S. (1976). Increasing attitude-behavior correspondence by broadening the scope of the behavioral measure. Journal of Personality and Social Psychology, 33, 793-802.

Weigel, R. H., Vernon, D. T. A., & Tognacci, L. N. (1974). Specificity of attitude as a determinant of attitude-behavior congruence. Journal of Personality and Social Psychology, 30, 724-728.

1

2

APPENDIX B: REFERENCES

(2005). "Merriam Webster Online Dictionary." from http://www.m-w.com/cgi-bin/dictionary.

Aldred, J. and M. Jacobs (2000). "Citizens and Wetlands: Evaluating the Ely Citizens' Jury." Ecological Economics 34: 217-232.

Alvarez-Farizo, B. and N. Hanley (2006). "Improving the Process of Valuing Non-Market Benefits: Combining Citizens' Juries with Choice Modelling." Land Economics 82(3): 465-478.

Andreoni, J. 1989. Giving with impure altruism: applications to charity and Ricardian equivalence. Journal of Political Economy 97:1447-1458.

———. 1990. Impure altruism and donations to public goods: a theory of warm-glow giving. Economic Journal 100:464-477.

Arrow, Kenneth J., Maureen L. Cropper, George C. Eads, Robert W. Hahn, Lester B. Lave, Roger G. Noll, Paul R. Portney, Milton Russell, Richard Schmalensee, V. Kerry Smith, and Robert N. Stavins. 1996. "Is There a Role for Benefit-Cost Analysis in Environmental, Health, and Safety Regulation?" Science, 272 (April 12): 221-222.

Arrow, K., R. Solow, P.R. Portney, E.E. Leamer, R. Radner, and H. Schuman. 1993. Report of the NOAA Panel on Contingent Valuation. Washington, D.C: Government Printing Office.

Arvai, J. L. 2003. Using risk communication to disclose the outcome of a participatory decision making process: Effects on the perceived acceptability of risk-policy decisions. Risk Analysis 23:281-289.

Arvai, J. L., and R. Gregory. 2003. A decision focused approach for identifying cleanup priorities at contaminated sites. Environmental Science & Technology 37:1469-1476.

Arvai, Joseph, and Robin Gregory. 2003. Testing Alternative Decision Approaches for Identifying Cleanup Priorities at Contaminated Sites. Environmental Science and Technology 37:1469 -1476.

Arvai, J. L., R. Gregory, and T. McDaniels. 2001. Testing a structured decision approach: Value-focused thinking for deliberative risk communication. Risk Analysis 21:1065-1076.

Arvai, J. L., T. McDaniels, and R. Gregory. 2002. Exploring a structured decision approach for fostering participatory space policy making at NASA. Space Policy 18:221-231.

Ashenfelter, Orley, and Michael Greenstone. 2004. Using Mandated Speed Limits to Measure the Value of a Statistical Life. Journal of Political Economy 112:S226-S267.

Ascher, William. 1978. Forecasting: An Appraisal for Policymakers and Planners. Baltimore: Johns Hopkins University Press.

Ascher, William, and William H. Overholt. 1984. Strategic Planning and Forecasting. New York: John Wiley & Sons.

Assessment, Millennium Ecosystem. 2005. Ecosystems and Human Well-being: Synthesis. Washington, DC: Island Press.

Ayyub, Bilal M. 2001. Expert Elicitation of Expert Opinions for Uncertainty and Risks Boca Raton, FL: CRC Press.

Barbier, E.B. 2001. A note on the economics of biological invasions. Ecological Economics 39:197-202.

Beierle, T. C. 2002. The quality of stakeholder-based decisions. Risk Analysis 22:739-749.

Beierle, T., and J. Cayford. 2002. Democracy in Practice: Public Participation in Environmental Decisions. Resources for the Future.

Binder, C., R. M. Boumans, R, Costanza. 2003. Applying the Patuxent Landscape Unit Model to human dominated ecosystems: the case of agriculture. Ecological Modelling 159.

Blamey, R.K., et al. (2000). Citizens' Juries and Environmental Value Assessment. Canberra, Australia, Research School of Social Sciences, Australian National University.

Bockstael, N., R. Costanza, et al. (1995). "Ecological economic modeling and valuation of ecosystems." Ecological Economics 14: 143-159.

Bockstael, N., R. Costanza, I. Strand, W. Boynton, K. Bell, and L. Wainger. 1995. Ecological economic modeling and valuation of ecosystems. Ecological Economics 14:143-159.

Bockstael, N.E., A.Myrick Freeman, R. Kopp, P.R. Portney, and V.K. Smith. 2000. On measuring economic values for nature. Environmental Science and Technology 34:1384-1389.

Binder, C., R. M. Boumans, R, Costanza (2003). "Applying the Patuxent Landscape Unit Model to human dominated ecosystems: the case of agriculture." Ecological Modelling 159.

Blamey, R.K., RF James, R Smith and S Niemeyer. 2000. Citizens' Juries and Environmental Value Assessment. http://cjp.anu.edu.au/docs/CJ1.pdf.: The first in a series of reports to be published containing the results of the research project Citizens' Juries for Environmental Management

Bond, P., and P. Goldblatt. 1984. Plants of the Cape Flora. Journal of South African Botany Suppl. 13:1-455.

Bongers, T. and H. Ferris.1999. Nematode community structure as a bioindicator in environmental monitoring. TREE 14: 224-228

Boyd, James. 2004. What's Nature Worth? Using Indicators to Open the Black Box of Ecological Valuation. Resources.

Boyd, James, and Spencer Banzhaf. 2006. What are Ecosystem Services? The Need for Standardized Environmental Accounting Units. In RFF Discussion Papers. Washington, DC: Resources for the Future.

Brink, D. O. (1989). Moral Realism and the Foundation of Ethics. Cambridge, Cambridge University Press.

Brouwer, Roy, and Ian J. Bateman. 2005. Benefits Transfer of Willingness to Pay Estimates and Functions for Health-Risk Reductions: A Cross-Country Study.

Journal of Health Economics 24:591-611.

Budescu DF, Wallsten TS. Processing linguistic probabilities: general principles and empirical evidence. In: Busemeyer JR, Hastie R, Medin DL, eds. Decision making from the perspective of cognitive psychology. New York: Academic Press, 1995:275-318.

Budnitz, R.J., G. Apostolakis, D.M. Boore, L.S. Cluff, K.J. Coppersmith, C.A. Cornell, & P.A. Morris. 1997. Recommendations for Probabilistic Seismic Hazard Analysis: Guidance on Uncertainty and the Use of Experts. Senior Seismic Hazard Analysis Committee. Washington, D.C.: U.S. Nuclear Regulatory Commission.

Carpenter, S., W. Brock, and P. Hanson. 1999. Ecological and social dynamics in simple models of ecosystem management. Conservation Ecology 3(2): 4. [online] URL: http://www.consecol.org/vol3/iss2/art4/

Carpenter ,S.R., F. Caraco, L. Correll,, W. Howarth, N. Sharpley, and V. H. Smith. 1998. Nonpoint pollution of surface waters with phosphorus and nitrogen. Ecological Applications 8: 559-568.

Carpenter, S.R., R.DeFries, T.Dietz, H.A. Mooney, S. Polasky, W.V. Reid and R.J.Scholes. 2006. Millennium Ecosystem Assessment: Research Needs. Science 314:257-258.

Carson, R.T., T. Groves, and M. Machina. 1999. Incentive and Informational Properties of Preference Questions: University of California, San Diego.

Carson, R.T., and R.C. Mitchell. 1993. The Value of Clean Water: The Public's Willingness to Pay for Boatable, Fishable, and Swimmable Quality Water. Water Resources Research 29 (July):2445-54.

Checkland, P. 1989. Soft Systems Methodology. In Rational Analysis for a Problematic World, edited by J. Rosenhead. Chichester, England: John Wiley and Sons.

Chess, C., and K. Purcell. 1999. Public participation and the environment: Do we know what works? Environmental Science & Technology 33:2685-2691.

Chestnut, L.G., and R.D. Rowe. 1990. Preservation Value for Visibility Protection at

the National Parks: U.S. Environmental Protection Agency.

Cleaves, D.A. 1994. Assessing Uncertainty in Expert Judgments about Natural Resources. New Orleans: U.S. Forest Service General Technical Report SO-1, July.

Clemen, R. T. 1996. Making Hard Decisions: An Introduction to Decision Analysis. PWS-Kent Publishing Co., Boston, MA.

Cooke, Roger M. 1991. Experts in Uncertainty: Expert Opinion and Subjective Probability in Science. Oxford: Oxford University Press.

Costanza, R. 1987. Simulation Modeling on the Macintosh Using STELLA. BioScience 37:129 - 132.

Costanza, R. (2004). Value Theory and Energy. Encyclopedia of Energy.

Costanza, Robert, and Steven Farber. 1986. The Economic Value of Coastal Louisiana Wetlands. Baton Rouge: Louisiana State University.

Costanza, Robert, and Ruth Matthias. 1998. Using Dynamic Modeling to Scope Environmental Problems and Build Consensus. Environmental Management 22:183-195.

Costanza, R., F. H. Sklar, and M. L. White. 1990. Modeling coastal landscape dynamics. BioScience 40:91-107.

Costanza, R., and A. Voinov, eds. 2003. Landscape Simulation Modeling: A Spatially Explicit, Dynamic Approach. New York: Springer.

Costanza, R., A. Voinov, R. Boumans, T. Maxwell, F. Villa, L. Wainger, H. Voinov. 2002. Integrated ecological economic modeling of the Patuxent River watershed, Maryland. Ecological Monographs 72:201-231.

Costanza, R., L. Wainger, C. Folke, and K-G Mäler. 1993. Modeling complex ecological economic systems: toward an evolutionary, dynamic understanding of people and nature. BioScience 43:545-555.

Covich, A.P., K.C. Ewel, R.O. Hall, P. S. Giller, W. Goedkoop, and D.M. Merritt. 2004. Ecosystem services provided by freshwater benthos. pp 45-72 In. Sustaining Biodiversity and Ecosystem Services in Soils and Sediments. D.H.

Wall Editor. Island Press, Washington, DC

Cowling, R., and R. Costanza. 1997. Valuation and Management of Fynbos Ecosystems. Ecological Economics 22:103-155.

Cowling, R.M. 1992. The ecology of fynbos. Nutrients, fire and diversity. Cape Town: Oxford University Press.

Davis, Elaine F. 2001. Reuse Assessment: A Tool to Implement The Superfund Land Use Directive. Memorandum to Superfund National Policy Managers. In, edited by O.-P. U.S. Environmental Protection Agency, Signed June 4, 2001. 24 pp. OSWER 9355.7-06P, Signed June 4, 2001.

de Zwart, D., S. D. Dyer, L. Posthuma, and C. P. Hawkins. 2006. Predictive models attributer effects on fish assemblages to toxicity and habitat alteration. Ecological Applications 16(4)1295–1310.

Deacon, R., and P. Shapiro. 1975. Private preference for collective goods revealed through voting on referenda. American Economic Review 65:793.

Desvouges, W.H., F.R. Johnson, R.W. Dunford, K.J. Boyle, S.P. Hudson, and K.N. Wilson. 1993. Measuring natural resource damages with contingent valuation: tests of validity and reliability. In Contingent Valuation: A Critical Assessment, edited by J. Hausman. Amsterdam: North Holland Press.

Edwards, S.F. 1986. Ethical perspectives and the assessment of existence values: does the neoclassical model fit? Northeast Journal of Agric. Resour. Econ. 15:145-159.

———. 1992. Rethinking existence values. Land Economics 68:120- 122.

Efroymson, R.A., J.P. Nicolette, and G.W. Suter. 2004. A Framework for Net Environmental Benefits Analysis for Remediation or Restoration of Contaminated Sites. Environmental Management 34 (3):315-331.

Egan, Kevin. 2004. Recreation Demand: On-site Sampling and Responsiveness of Trip Behavior to Physical Water Quality Measures. Unpublished PhD. Thesis. Iowa State University.

Egan, Kevin, and Herriges, Joseph. 2006. Multivariate Count Data Regression

Models with Individual Panel Data from and On-site Sample. Journal of Environmental Economics and Management 52:567-581.

Egan, Kevin, and Herriges, Joseph. 2006. Multivariate Count Data Regression Models with Individual Panel Data from and On-site Sample. Journal of Environmental Economics and Management 52:567-581.

Encarnacao, J., Foley, J., Bryson, S., Feiner, S. K., & Gershon, N. (1994). Research issues in perception and user interfaces. Computer Graphics and Applications, IEEE, 14(2), 67-69.

Failling, L. and R. Gregory, 2003. Ten common mistakes in designing biodiversity indicators for forest policy. Journal of Environmental Management Volume 68, Issue 2 , Pages 121-132.

Farber, Stephen, Robert Costanza, Daniel L. Childers, Jon Erickson, Katherine Gross, Morgan Grove, Charles S. Hopkinson, James Kahn, Stephanie Pincetl, Austin Troy, Paige Warren, and Matthew Wilson. 2006. Linking Ecology and Economics for Ecosystem Management. BioScience:117-129.

Failing, L., and R. Gregory. 2003. Ten common mistakes in designing biodiversity indicators for forest policy. Journal of Environmental Management 68 (2):121-132.

Farber, Stephen, et al. 2006. Linking Ecology and Economics for Ecosystem Management," BioScience 56:117-129.

Fischhoff, B. 1997. What do psychologists want? Contingent valuation as a special case of asking questions. In Determining the Value of Non-Marketed Goods, edited by R. J. Kopp, W. W. Pommerehne and N. Schwarz. Norwell, MA: Kluwer Academic Publishers.

Fitz, H. C., E. B. DeBellevue, R. Costanza, R. Boumans, T. Maxwell, L. Wainger, and F. H. Sklar. 1996. 1996. Development of a general ecosystem model (GEM) for a range of scales and ecosystems. Ecological Modelling 88:263-297.

Ford, A. 1999. Modeling the Environment: An Introduction to System Dynamics Models of Environmental Systems. Washington, DC: Island Press.

Freeman, A. Myrick III. 1997. The Benefits of Water Quality Improvements for Marine Recreation: A Review of the Empirical Evidence. Marine Resource Economics 10:385-406.

Freeman, A.M. III. 2003. The Measurement of Environmental and Resource Values: Theory and Methods. 2nd Edition ed. Washington, D.C.: Resources for the Future.

Gelso, Brett R., and Jeffrey M. Peterson. 2005. The influence of ethical attitudes on the demand for environmental recreation: incorporating lexicographic preferences. Ecological Economics 53:35-45.

Glover, J. (1984). What Sort of People Should There Be? Penguin.

Goodpaster, K. (1978). "On Being Morally Considerable." Journal of Philosophy 75: 308-25.

Gregory, R. 2000. Using stakeholder values to make smarter environmental decisions. Environmental and Resource Economics 42:34-44.

Gregory, R., J. L. Arvai, and T. McDaniels. 2001a. Value-focused thinking for environmental risk consultations. Research in Social Problems and Public Policy 9:249-275.

Gregory, R., T. McDaniels, and D. Fields. 2001b. Decision aiding, not dispute resolution: Creating insights through structured environmental decisions. Journal of Policy Analysis and Management 20:415-432.

Gregory, Robin, Sarah Lichtenstein, and Paul Slovic. 1993. Valuing Environmental Resources; A Constructive Approach. Journal of Risk and Uncertainty 7:177-197.

Gregory, Robin, and Katharine Wellman. 2001. Bringing stakeholder values into environmental policy choices: a community-based estuary case study. Ecological Economics 39:37-52.

Groombridge, Brian, and Martin Jenkins. 1998. “Freshwater Biodiversity: A Preliminary Global Assessment,” WCMC Biodiversity Series No. 8. Cambridge: World Conservation Press.

Grossman, D.H., and P.J. Comer. 2004. Setting Priorities for Biodiversity Conservation in Puerto Rico: NatureServe Technical Report.

Gunderson, L. C., S. Holling, and S. Light. 1995. Barriers and bridges to the renewal of ecosystems and institutions. New York: Columbia University Press.

Hammond, J., R. L. Keeney, and H. Raiffa. 1999. Smart Choices: A Practical Guide to Making Better Decisions. Cambridge, MA: Harvard Business School Press.

Hannon, B., and M. Ruth. 1997. Modeling Dynamic Biological Systems. New York: Springer-Verlag.

Hannon, B., and M. Ruth. 1994. Dynamic Modeling. New York: Springer-Verlag.

Harrison, G. W. and E. E. Rutström (1999). Experimental Evidence on the Existence of Hypothetical Bias in Value Elicitation Methods. Handbook of Results in Experimental Economics. V. L. Smith. New York, Elsevier Science.

Harwell, M. V. Myers, T. Young, A. Bartuska, N. Gassman, J.H.Gentile, C. Harwell, S. Appelbaum, J. Barko, B.Causey, C. Johnson, A. McLean, R. Smola, P. Templet, and S. Tosini. 1999. A framework for an ecosystem integrity report card. BioScience 49(7): 543-554.

Hays, S. P. (1989). Beauty, Health and Permanence: Environmental Politics in the United States, 1955-1985, Cambridge.

Helmer, Olaf. 1967. Analysis of the Future: The Delphi Method. Report P-3558. RAND Corporation. Santa Monica, CA.

Higgins, S. I., J. K. Turpie, R. Costanza, R. M. Cowling, D. C. le Maitre, C. Marais, and G. Midgley. 1997. An ecological economic simulation model of mountain fynbos ecosystems: dynamics, valuation, and management. Ecological Economics 22:155-169.

Hitlin, S. and Piliavin, J. A., 2004. Values: Reviving a Dormant Concept. Annual Review of Sociology, 30:359–93.

Hoagland, P. and D. Jin. 2006. Science and economics in the management of invasive species. BioScience 56(11):931-935.

Holling, C. S. 1999. Introduction to the special feature: just complex enough for

understanding; just simple enough for communication. Conservation Ecology 3(2): 1. [online] URL: http://www.consecol.org/vol3/iss2/art1/

Horowitz, J.A. and K.E. McConnell, 2002. A Review of WTA / WTP Studies. Journal of Environmental Economics and Management 44, 426 -447.

Hypothetical Bias in Value Elicitation Methods. In Handbook of Results in Experimental Economics, edited by V. L. Smith. New York: Elsevier Science.

Hays, Samuel P. 1989. Beauty, Health and Permanence: Environmental Politics in the United States, 1955-1985: Cambridge.

Heiman, M. 1990. From “not in my backyard!” to “not in anybody’s backyard!” Journal of the American Planning Association 56:359-362.

Helmer, Olaf. 1967. Analysis of the Future: The Delphi Method. Santa Monica, CA.: RAND Corporation.

Higgins, S. I., J. K. Turpie, et al. (1997). "An ecological economic simulation model of mountain fynbos ecosystems: dynamics, valuation, and management." Ecological Economics 22: 155-169.

Higgins, S. I., J. K. Turpie, R. Costanza, R. M. Cowling, D. C. le Maitre, C. Marais, and G. Midgley. 1997. An ecological economic simulation model of mountain fynbos ecosystems: dynamics, valuation, and management. Ecological Economics 22:155-169.

Hillier, J. 1998. Beyond confused noise: Ideas toward communicative procedural justice. Journal of Planning Education and Research 18:14-24.

Hobbs, R.J., D.M. Richardson, and G.W. Davis. 1995. Mediterranean-type ecosystems: opportunities and constraints for studying the function of biodiversity. In Mediterranean-type ecosystems. The function of biodiversity., edited by G. W. Davis and D. M. Richardson. Berlin: Springer.

Interstate Technology Regulatory Council. 2006. Planning and Promoting Ecological Land Reuse of Remediated Sites: (www. itrcweb.org).

Involvement in Local Water Management Decisions. Risk Analysis 19. 9??(Full citation?)

Jaffe, Judson, and Robert N. Stavins. 2004. The Value of Formal Quantitative Assessment of Uncertainty in Regulatory Analysis. AEI-Brookings Jointed Center for Regulatory Studies, Related Publication 04-22.

James, R.F. & R.K. Blamey (2000). A Citizens' Jury Study of National Park Management. Canberra, Australia, Research School of Social Sciences, Australian National University.

Janis, Irving. 1982. Groupthink: Psychological Studies of Policy Decisions and Fiascoes. Second Edition. New York: Houghton Mifflin.

Janssen, M. A. and S. R. Carpenter. 1999. Managing the Resilience of Lakes: A multi-agent modeling approach. Conservation Ecology 3(2): 15. [online] URL: http://www.consecol.org/vol3/iss2/art15/

Jenkins, Robin R., Nicole Owens, and Lanelle Bembenek Wiggins. 2001. Valuing Reduced Risks to Children: The Case of Bicycle Safety Helmets. Contemporary Economic Policy 19:4, 397-408.

Johnston, R.J., E.Y. Besedin, and R.F. Wordwell. 2003. Modeling Relationships Between Use and Nonuse Values for Surface Water Quality: A Meta-Analysis. Water Resources Research 39 (12):1363-1372.

Kahn, M.E., and J.G. Matsusaka. Demand for environmental goods: Evidence from voting patterns on California initiatives. Journal of Law and Economics 40:137-173.

———. 1997. Demand for environmental goods: Evidence from voting patterns on California initiatives. Journal of Law and Economics 40:137-173.

Kahnemann, D., P. Slovic, and A. Tversky. 1982. Judgment Under Uncertainty: Heuristics and Biasis. Cambridge: Cambridge University Press.

Kahnemann, D., and A. Tversky. 1974. Judgment Under Uncertainty. Science 185:1124 - 1131.

Karr, J.R. 1981. Assessment of biotic integrity using fish communities. Fisheries 6(6):21-27.

Karr, J.R. 1993. Defining and assessing ecological integrity: beyond water quality.

Environmental Toxicology and Chhemistry 12:15621-1531.

Keeney, R. L., and R. Gregory. 2005. Selecting attributes to measure the achievement of objectives. Operations Research 53:1-11.

Keeney, R. L., and H. Raiffa. 1993. Decisions with multiple objectives: Preferences and value tradeoffs. Cambridge University Press, Cambridge, UK.

Keeney, R., D. von Winterfeldt, and T. Eppel. 1990. Eliciting public values for complex policy decisions. Management Science 36:1011-1030.

Kenyon, W. et al. (2001). "Citizens' Juries: An Aid to Environmental Valuation?" Environmental Planning C: Government & Policy 19(4): 557-566.

Korsgaard, C. (1996). Two Distinctions in Goodness. Creating the Kingdom of Ends. C. Korsgaard. Cambridge, Cambridge University Press. 1996: 249-74.

Kraft, M. E. 1988. Evaluating technology through public participation: The nuclear waste disposal controversy. In Technology and Politics, edited by M. E. Kraft and N. J. Vig. Durham, NC: Duke University Press.

Kraft, M. E., and D. Scheberle. 1995. Environmental justice and the allocation of risk: The case of lead and public health. Policy Studies Journal 23:113-122.

Kruse, S. 2005. Creating an interdisciplinary framework for ecopnomic valuation: A CVM approach to dam removal, Department of Agricultural, Environmental, and Developmental Economics, The Ohio State University, Columbus, OH.

Larkin,R. 2006 Effect of swine and dairy manure amendments on microbial communities in three soils as influenced by environmental conditions. Biology & Fertility of Soils. 43:51

Leung, B., D. Finnoff, J.F. Shogren. and D. Lodge. 2005. Managing invasive species: Rules of thumb for rapid assessment. Ecological Economics 55:24-36.

Levitan, L., I. Merwin, and J. Kovach. 1995. "Assessing the Relative Environmental Impacts Of Agricultural Pesticides: The Quest for a Holistic Method", Agriculture, Ecosystems and Environment. 55: 153-168.

Lichtenstein, Sarah, Baruch Fischhoff, and Lawrence D. Phillips, "Calibration of

Probabilities: The State of the Art to 1980," ," in Daniel Kahneman, Paul Slovic, and Amos Tversky, eds., Judgment under Uncertainty: Heuristics and Biases. New York: Cambridge University Press, pp. 306-334.

Lind, E., and T. Tyler. 1988. The Social Psychology of Procedural Justice. New York, NY: Plenum Press.

List, J., and C. Gallet. 2001. What Experimental Protocol Influence Disparities Between Actual and Hypothetical Stated Values? Evidence from a Meta-Analysis. Environmental and Resource Economics 20:241-254.

Loomis, J., T., B. Lucero Brown, and G. Peterson. 1996. Improving Validity Experiments of Contingent Valuation Methods: Results of Efforts to Reduce the Disparity of Hypothetical and Actual Willingness to Pay. Land Economics 72:4450-4461.

Luce, M. F. 1998. Choosing to avoid: Coping with negatively emotion-laden consumer decisions. Journal of Consumer Research 24:409-433.

Lyneis, J.M. 1980. Corporate Planning and Policy Design: A System Dynamics Approach. Cambridge, Massachusetts: Pugh-Roberts Associates.

McDaniels, T., R. Gregory, J. L. Arvai, and R. Chuenpagdee. 2003. Decision structuring as a means of alleviating embedding in environmental valuation. Ecological Economics.

McDaniels, Timothy L., Robin S. Gregory, and Daryl Fields. 1999. Democratizing Risk Management: Successful Public

McDowell, J. (1985). Values and Secondary Qualities. Morality and Objectivity. T. Honderich, Routledge and Kegan Paul: 110-29.

MacEachren, A. M. (1995). How Maps Work. New York and London: The Guilford Press.

MacEachren, A. M., Robinso, A., Hopper, S., Gardner, S., Murrary, R., Gahegan, M., et al. (2005). Visualizing geospatial information uncertainty: what we know and what we need to know. Cartography and Geographic Information Science, 32(3), 139 - 160.

Macmillan, D.C., et al. (2002). "Valuing the Non-Market Benefits of Wild Goose Conservation: A Comparison of Interview and Group-Based Approaches." Ecological Economics 43: 49-59.

Mahan, B.L., S. Polasky, and R.M. Adams, 2000, "Valuing Urban Wetlands: A Property Price Approach," Land Economics, 76 (February): 100-113.

Merriam Webster Online Dictionary 2005 [cited. Available from http://www.m-w.com/cgi-bin/dictionary.

Millennium Ecosystem Assessment Board. 2003. Ecosystems and Human Well-being; A Report of the Conceptual Framework Working Group of the Millennium Ecosystem Assessment. Washington, DC: Island Press.

Mitchell, R., and R. Carson. 1995. Current issues in the design, administration, and analysis of contingent valuation surveys. In Current issues in environmental economics, edited by P. Johansson, B. Kristrom and K. Maler. New York, NY: Manchester University Press.

Montgomery, M., and Needelman, M. 1997. The Welfare Effects of Toxic Contamination in Freshwater Fish. Land Economics 73:211-223.

Morecroft, J.D.W., D.C. Lane, and P.S. Viita. 1991. Modelling Growth Strategy in a Biotechnology Startup Firm. System Dynamics Review 7:93-116.

Morecroft, J.D.W., and J.D. Sterman. 1994. Modeling for Learning Organizations. Portland, Oregon: Productivity Press.

Murphy, James J., P. Geoffrey Allen, Thomas H. Stevens, & Darryl Weatherhead. 2003. A Meta-Analysis of Hypothetical Bias in Stated Preference Valuation. University of Massachusetts, Amherst Working Paper No. 2003-8.

Myers, N. 1990. The biodiversity challenge: expanded hot-spots analysis. The Environmentalist 10:243-255.

Nachman, K.E., Graham, J.P., Price, L., Sibergeld, E. 2005. Arsenic: A roadblock to potential animal waste management solutions. Environmental Health Perspectives 113: 1123-1124

National Research Council. 1989. Improving Risk Communication. Washington, DC:

National Academy Press.

———. 1996. Understanding Risk: Informing Decisions in a Democratic Society. National Academy Press.

———.1997. Review of Recommendations for Probabilistic Seismic Hazard Analysis: Guidance on Uncertainty and Use of Experts. Washington: National Academies Press.

———. 2001. The Science of Regional and Global Change: Putting Knowledge to Work.

———. 2002. Estimating the Public Health Benefits of Proposed Air Pollution Regulations. The National Academies Press: Washington, D.C.

———. 2003. Air Emissions from Animal Feeding Operations: Current Knowledge, Future Needs. Washington, DC

———.2004. Valuing ecosystem services; toward better environmental decision-making. National Academy Press, Washington, DC.

Navrud, Stale, and Richard Ready, eds., Environmental Value Transfer: Issues and Methods, Springer: Dordrecht, The Netherlands, in press.

Netusil, Noelwah, 2005, "The Effect of Environmental Zoning and Amenities on Property Values: Portland Oregon," Land Economics, 81 (May): 227.

Niemi GJ, McDonald ME. Application of ecological indicators . Annual Review of Ecology Evolution and Systematics 35: 89-111 2004.

Opaluch, James, Thomas Grigalunas, Jerry Diamantides, Marisa Mazzotta, and Robert Johnston. 1999. Recreational and Resource Economic Values for the Peconic Estuary System. Riverhead, N.Y.: Peconic Estuary Program, Suffolk County Department of Health Services County Center.

Oskamp, Stuart. 1982. "Over-Confidence in Case-Study Judgment," in Daniel Kahneman, Paul Slovic, and Amos Tversky, eds., Judgment under Uncertainty: Heuristics and Biases. New York: Cambridge University Press, pp. 287-293

Oster, G. Madonna 1996 [cited. Available from

http://nature.berkeley.edu/~goster/madonna.html.

Palmquist, Raymond B. 2005. Hedonic Models. In Handbook of Environmental Economics, edited by K. Mäler and J. Vincent: North Holland.

Payne, J. W., J. Bettman, R.Johnson, and E. J. Johnson. 1993. The Adaptive Decision Maker. Cambridge, MA: Cambridge University Press.

Peterson, S. 1994. Software for Model Building and Simulation: An Illustration of Design Philosophy. In Modeling for Learning Organizations, edited by J. D. W. Morecroft and J. D. Sterman. Productivity Press: Portland, Oregon.

Phaneuf, Daniel. 2002. A Random Utility Model for TMDLs: Estimating the benefits of Watershed Based Ambient Water Quality Improvements. Water Resources Research 38:366.1-36-11.

Paheuf, Daniel, Palmquist, Raymond, and Smith, V. Kerry, 2006. **Full citation??**

Phaneuf, Daniel and V. Kerry Smith, 2005, "Recreation Demand Models" in K.G. Maler and J. R. Vincent , editors, Handbook of Environmental Economics , vol 2 (North Holland).

Phillips, L.D. 1990. Decision Analysis for Group Decision Support. In Tackling Strategic Problems: The Role of Group Decision Support, edited by C. Eden and J. Radford. London: Sage Publishers.

Polasky, S., O. Gainutdinova, and J. Kerkvliet. 1996. Comparing CV Responses with Voting Behavior: Open Space Survey and Referendum in Corvallis Oregon. Paper read at Annual U.S.D.A. W-133 meeting, February 1996, at Jekyll Island, GA.

Polasky, Stephen, Christopher Costello, and Andrew Solow. In press. The Economics of Biodiversity. In The Handbook of Environmental Economics, edited by J. Vincent and K. Maler.

Portney, P. R. 1994. The contingent valuation debate: Why economists should care. Journal of Economic Perspectives 8:3-17.

Poulos, Christine, V. Kerry Smith, and Hyun Kim, 2002, "Treating Open Space as an Urban Amenity," Resource and Energy Economics 24: 107-129.

Rao, S.S. 1994. Welcome to open space. Training (April):52-55.

Rempel RS, Andison DW, Hannon SJ Guiding principles for developing an indicator and monitoring framework. Foresty Chronicle 80 (1): 82-90 JAN-FEB 2004

Renn, O. 1999. A model for analytic-deliberative process in risk management. Environmental Science & Technology 33:3049-3055.

Richmond, B., and S. Peterson. 1994. STELLA II Documentation. Hanover, New Hampshire: High Performance Systems, Inc.

Roberts, E.B. 1978. Managerial Applications of System Dynamics. Portland, Oregon: Productivity Press.

Rolston III, H. (1991). Environmental Ethics: Values in and Duties to the Natural World. The Broken Circle: Ecology, Economics, and Ethics. F. H. B. a. S. Kellert. New Haven, Yale University Press.

Romney, A. K., S. C. Weller, and W. H. Batchelder. 1986. Culture as Consensus: A theory of culture and informant accuracy. Am. Anthropol. 88:313-338.

Romney, A. Kimball, Carmella C. Moore, William H. Batchelder, and Ti-Lien Hsia. 2000. Statistical methods for characterizing similarities and differences between semantic structures. PNAS 97:518-523.

Rosenberger, Randall S., and Loomis, John B. 2003. Benefits Transfer. in Champ, Patricia A., Boyle, Kevin J., and Brown, Thomas C. A Primer on Nonmarket Valuation. Kluwer Academic Publishers, Dordrecht.

Rossi, Peter H., Howard Freeman and Mark W. Lipsey (1999). Evaluation, 6th edition. Thousand Oaks, CA: Sage Publications.

Roughgarden, Joan. 1998. Primer of Ecological Theory: Prentice Hall.

Roughgarden, J. 1998. Production functions from ecological populations: a survey with emphasis on spatially explicit models. In Spatial Ecology: The Rule of Space in Population Dynamics and Interspecific Interactions, edited by D. Tilman and P. Kareiva: Princeton University Press.

Rosenhead, J. 1989. Rational Analysis of a Problematic World. Chichester, England: John Wiley and Sons.

Russell III, Edmund P. 1993. Lost Among the Parts per Billion; Ecological Protection at the Environmental Protection Agency, 1970-1993. Environmental History 2:29-51.

Sackman, Harold. 1974. Delphi Assessment: Expert Opinion, Forecasting, and Group Processes. Santa Monica, CA: RAND Corporation.

Sagoff, Mark. 2004. Price, Principle, and the Environment. Cambridge: Cambridge University Press.

Sandman, P. M., Weinstein, N. D., & Miller, P. (1994). High-risk or low - how location on a risk ladder affects perceived risk. Risk. Risk Analysis, 14, 35 - 45.

Sayre-McCord, G. (1988). The Many Moral Realisms. Essays on Moral Realism. G. Sayre-McCord. Ithaca, Cornell University Press: 1-26.

Schiller, A., C. T. Hunsaker, M. A. Kane, A. K. Wolfe, V. H. Dale, G. W. Suter, C. S. Russell, G. Pion, M. H. Jensen, and V. C. Konar. 2001. Communicating ecological indicators to decision makers and the public. Conservation Ecology 5 (1):19. http://www.ecologyandsociety.org/vol5/iss1/art19/#FromValuesToValuedAspects

Schriver, Karen A, 1996. Dynamics in Document Design: Creating Text for Readers. Wiley.

Schwartz SH. 1994. Are there universal aspects in the structure and content of human values? J. Soc. Issues 50:19–45.

Scriven, M. (1967). The methodology of evaluation. In Perspectives of Curriculum Evaluation (American Educational Research Association Monograph Series on Curriculum Evaluation, No. 1). Chicago, IL: Rand McNally.

Senge, P.M. 1990. The Fifth Discipline. New York: Doubleday.

Settle, C., T.D. Crocker and J.F. Shogren. 2002. On the joint determination of biological and economic systems. Ecological Economics 42:301-311.

Shabman, L., and K. Stephenson. 1996. Searching for the correct benefit estimate:

Empirical evidence for an alternative perspective. Land Economics 72:433-49.

Shabman, Leonard A., and Sandra S. Batie. 1978. The Economic Value of Coastal Wetlands: A Critique. Coastal Zone Management Journal 4 (3):231-237.

Shah, P & Miyake, A. (Eds) The Cambridge Handbook of Visuospatial Thinking. Cambridge University Press, 2005.

Sidgwick, H. (1901). Methods of Ethics. New York, Macmillan.

Simon, H.A. 1956. Administrative Behavior. New York: Wiley and Sons.

———.1956. Rational choice and the structure of the environment. Psychological Review 63:129-138.

———. 1979. Rational Decision-Making in Business Organizations. American Economic Review 69:493 - 513.

Slovic, P. 1987. Perception of risk. Science 236:280-285.

———. 1995. The construction of preference. American Psychologist 50:364-371.

Slovic, Paul, Baruch Fischhoff, and Sarah Lichtenstein. 1982. "Facts vs. Fears: Understanding Perceived Risks," ," in Daniel Kahneman, Paul Slovic, and Amos Tversky, eds., Judgment under Uncertainty: Heuristics and Biases. New York: Cambridge University Press, pp. 463-492.

Smith, V. Kerry, and J.C. Huang. 1995. Can Markets Value Air Quality? A Meta-Analysis of Hedonic Property Value Models. Journal of Political Economy 103:209-27.

Smith, V. Kerry, and Yoshiaki Kaoru. 1990a. Signals or Noise? Explaining the Variation in Recreation Benefit Estimates. American Journal of Agricultural Economics 72 (2):419-433.

———. 1990b. What Have We Learned Since Hotelling's Letter? A Meta-Analysis. Economics Letters 32 (3): 267-272.

Smith, V. Kerry, and Subhrendu K. Pattanayak. 2002. Is Meta-Analysis a Noah's Ark for Non-Market Valuation? Environmental and Resource Economics 22 (1-2):271-296.

Spyridakis, J.H. Guidelines for Authoring Comprehensible Web Pages and

Evaluating Their Success. Technical Communication, 47, 3, 301-310, 2000, accessible at http://www.uwtc.washington.edu/people/faculty/jspyridakis.php)

Stoms, D. M., P. J. Comer, P. J. Crist, and D. H. Grossman. 2005. Choosing surrogates for biodiversity conservation in complex planning environments. *Journal of Conservation Planning* 1:44-63.

Strecher, V. J., Greenwood, T., Wang, C., & Dumont, D. (1999). Interactive Multimedia and Risk Communication. Journal of the National Cancer Institute. Monographs(25), 134-139.

Sturgeon, N. (1985). Moral Explanations. Morality, Reason, and Truth. D. C. a. D. Zimmerman, Rowman and Allenheld: 49-78.

Syme, G., D. Macpherson, and C. Seligman. 1991. Factors motivating community participation in regional water allocation planning. Environment and Planning A 23:1779-1795.

Taylor, P. (1986). Respect for Nature: A Theory of Environmental Ethics. Princeton, Princeton University Press.

Thorndike, E. L. A. 1920. Constant error in psychological ratings. Journal of Applied Psychology 4:25-29.

U.S. Environmental Protection Agency. 1984. Costs and Benefits of Reducing Lead in Gasoline – Draft Final Report, Washington, DC.

———. 1987. Unfinished Business: A Comparative Assessment of Environmental Problems - Overview Report: 230287025a.

———. 1994. Managing Ecological Risks at EPA: Issues and Recommendations for Progress. EPA/600/R-94/183.

———.. 1996. State of the Great Lakes.———. 1997. Guiding Principles for Monte Carlo Analysis. March. EPA/630/R-97/001.

———.1999. The Benefits and Costs of the Clean Air Act 1999 to 2010, Washington, DC.

———. 2000. Guidelines for Preparing Economic Analyses: EPA 240-R-00-003.

———. 2000. Engaging the American People: A Review of EPA's Public

Participation Policy and Regulations with Recommendations for Action. Washington, DC.

———.2001. Environmental Monitoring and Assessment Program: West—Research Strategy. (www.epa.gov/emap/wpilot/index.html)

———. 2002. Environmental and Economic Benefit Analysis of Final Revisions to the National Pollutant Discharge Elimination System (NPDES) Regulation and the Effluent Guidelines for Concentrated Animal Feeding Operations (CAFOs). EPA-821-R-03-003.

———. 2002. Draft Report on the Environment. Vol. EPA 260-R-02-006.

———. 2004. Innovating for Better Results: a Report on EPA Progress from the Innovations Action Council: EPA100-R-04-001

———.2004. Economic and Environmental Benefits Analysis of Final Effluent Limitations Guidelines and New Source Performance Standards for the Concentrated Aquatic Animal Production Industry Point Source Category, Washington, DC.

———. 2005. "Ecological Benefits Assessment Strategic Plan; SAB Review Draft, November 3, 2005."

———. 2006. Ecological Benefits Assessment Strategic Plan: Washington, DC., EPA-240-R-06-001.

U.S. Environmental Protection Agency CERCLA Education Center. 2005. Fundamentals of Superfund: Participant Manual.

U.S. Environmental Protection Agency Office of Policy, Economics and Innovation. 2003. Public Involvement Policy of the U.S. Environmental Protection Agency. EPA 233-B-03-002

U.S. Environmental Protection Agency Risk Assessment Forum. 1992. Framework for Ecological Risk Assessment: USEPA EPA/630/R-92/001.

———. 1998. Guidelines for Ecological Risk Assessment: EPA/630/R095/002F.

———. 2003. Generic Ecological Assessment Endpoints (GEAEs) for Ecological Risk Assessment. EPA/630/P-02/004F.

U.S. Environmental Protection Agency Science Advisory Board. 1990. Reducing Risk: Setting Priorities and Strategies for Environmental Protection: EPA-SAB-EC-90-021.

———. 2000. Toward Integrated Environmental Decision-making: EPA-SAB-EC-00-011.

———. 2001. Improved Science-Based Environmental Stakeholder Processes: An EPA Science Advisory Board Commentary. EPA-SAB-EC-COM-001-006.

———. 2003. Summary Minutes of the U.S. Environmental Protection Agency (EPA) Science Advisory Board (SAB) Committee on Valuing the Protection of Ecological Systems and Services Initial Background Workshop October 27, 2003, J. W. Marriott Hotel, Washington, DC.

———. 2004. Advisory on EPA's Draft Report on the Environment. EPA-SAB-05-004

U.S. Environmental Protection Agency Science Advisory Board Staff. 2004. Survey of Needs for Regions for Science-Based Information on the Value of Protecting Ecological Systems and Services . Background document for CVPESS Meeting Sept, 13-15, 2004. 63 pp.

U.S. Office of Management and Budget. 2003. Circular A-4. September 17.

van den Belt, M. 2004. Mediated Modeling: a systems dynamics approach to environmental consensus building: Island Press.

Vari, A., J. L. Mumpower, and P. Reagan-Ciricione. 1993. Low-Level Radioactive Waste Disposal Facility Siting Processes in the United States. Albany, NY: Center for Policy Research, State University of New York.

Vennix, J. A. M. 1996. Group Model Building: Facilitating Team Learning Using System Dynamics.

Vennix, J.A.M., and J.W. Gubbels. 1994. Knowledge Elicitation in Conceptual Model Building: A Case Study in Modeling a Regional Dutch Health Care System. In Modeling for Learning Organizations, edited by J. D. W. Morecroft and J. D. Sterman. Portland, Oregon: Productivity Press.

Von Haefen, R. H. 1999. Valuing Environmental Quality in a Repeated Discrete-Continuous Framework. Unpublished PhD. Thesis. Duke University.

Vossler, C.A., and J. Kerkvliet. 2003. A criterion validity test of the contingent valuation method: Comparing hypothetical and actual voting behavior for a public referendum. Journal of Environmental Economics and Management 45 (3):631-49.\

Wall, D.H. 2004. Sustaining Biodiversity and Ecosystem Services in Soils and Sediments. Island Press. Washington, DC

Weslawski, J.M. et al., 2004. Marine sedimetary biota as providers of ecosystem goods and services. Pgs 73-98. In. Wall, D.H. Sustaining Biodiversity and Ecosystem Services in Soils and Sediment. Island Press. Washington, DC

Westenholme, E.F. 1990. System Inquiry: A System Dynamics Approach. Chichester, England: John Wiley and Sons.

———. 1994. A Systematic Approach to Model Creation. In Modeling for Learning Organizations, edited by J. D. W. Morecroft and J. D. Sterman. Portland, Oregon: Productivity Press.

Weyant, J. et al. 1996. "Integrated Assessment of Climate Change: An Overview and Comparison of Approaches and Results," in Intergovernmental Panel on Climate Change (IPCC), Climate Change 1995: Economics and Social Dimensions of Climate Change: Scientific-Technical Analysis. Cambridge: Cambridge University Press, pp. 367-396.

Whittington, D. et al. (1992). "Giving Respondents Time to Think in Contingent Valuation Studies: A Developing Country Example." Journal of Environmental Economics and Management 22: 205-22

Williams, B. (1994). Must a Concern for the Environment Be Centred on Human Beings? Reflecting on Nature. L. Gruen and D. Jamieson. Oxford, Oxford University Press.

Wilson, M. A. (2004). Ecosystem Services at Superfund Redevelopment Sites, Revealing the Value of Revitalized Landscapes through the Integration of Ecology and Economics. Report prepared by Spatial Informatics Group,

LLC under subcontract of Systems Research and Applications Corporation. Report funded by EPA's funded by EPA's Office of Solid Waste and Emergency Response (OSWER).

Worm, B., E.B. Barbier, N. Beaumont, J.E. Duffy, C. Folke, B.S. Halpern, J.B.C. Jackson, H.K. Lotze, F. Micheli, S.R. Palumbi, E. Sala, K.A. Selkoe, J.J. Stachowicz and R. Watson. 2006. Impacts of biodiversity loss on ocean ecosystem services. Science 314:787-790.

Young, T.F. and S. Sanzone (eds.).2002. A framework for assessing and reporting on ecological condition. EPA Science Advisory Board. U.S. Environmental Protection Agency. Washington, DC. (http://www.epa.gov/sab/pdf/epec02009.pdf) (Check/correct format for this citation)

1 # APPENDIX C: ENDNOTES

[1]Laws include: the Clean Air Act, Clean Water Act, Comprehensive Environmental Response, Compensation, and Liability Act, Federal Insecticide, Fungicide and Rodenticide Act, Toxic Substances Control Act, and Resource Conservation and Recovery Act

[2] Although CVPESS was initiated by the SAB, Senior EPA managers supported the concept of this SAB project and participated in the initial background workshop that launched the work of the C-VPESS.

[3] The SAB Staff Office published a Federal Register Notice on March 7, 2003 (68 FR 11082-11084) announcing the project and called for the public to nominate experts in the following areas: decision science; ecology; economics; engineering; psychology; and social sciences with emphasis in ecosystem protection. The SAB Staff Office published a memorandum on August 11, 2003 documenting the steps involved in forming the new committee and finalizing its membership.

[4] The committee developed the conclusions in this report after multiple public meetings and workshops: a) an Initial Background Workshop on October 27, 2003 to learn the range of EPA's needs for science-based information on valuing the protection of ecological systems and services from managers of EPA Headquarters and Regional Offices; b) a Workshop on Different Approaches and Methods for Valuing the Protection of Ecological Systems and Services, held on April 13-14, 2004; c) an advisory meeting focused on support documents for national rulemakings held on June 14-15, 2004; d) an advisory meeting focused on regional science needs, in EPA's Region 9 (San Francisco) Office on Sept. 13, 14, and 15, 2004; e) advisory meetings held on January 26-26, 2005 and April 12-13, 2005 to review EPA's draft *Ecological Benefits Assessment Strategic Plan;*and f) a Workshop on Science for Valuation of EPA's Ecological Protection Decisions and Programs, held on December 13-14, 2005 to discuss the integrated and expanded approach described in this paper. The also committee discussed text drafted for this report at public meetings on October 25 2005; May 9, 2006; October 5-6, 2006, and **(insert additional dates).**

[5] It is important to acknowledge that they are many different specific definitions for the term "benefits".
When used in technical discussions, different definitions make different assumptions that can be important to how the concept is interpreted. This implies that not all technical concepts for benefits are equivalent. Throughout this report, we use the term "benefits" to refer to the general concept of contributions to people. Economists generally use the term to refer more specifically to compensating or equivalent variations as measures of changes in the well-being of individuals (see Freeman, 2003). Despite these differences, the term services and the listing of types of services here refers exclusively to the contributions of ecosystems to human well being. As such, it is based on an anthropocentric concept of the benefits provided by ecosystems.

[6] Even without any subsequent valuation, explicitly listing the services derived from an ecosystem, and using the best available methods in the ecological, social, and behavioral sciences to help develop that list, can ensure appropriate recognition of the full range of

potential impacts of a given policy option. This can help make the analysis of ecological science more transparent and accessible and can help inform decision makers of the relative merits of different options before them.

[7] There is controversy over the meaning of intrinsic value (Korsgaard, C. (1996). Two Distinctions in Goodness. Creating the Kingdom of Ends. C. Korsgaard. Cambridge, Cambridge University Press. **1996:** 249-74. Many people take intrinsic value to mean that the value of something is inherent in that thing. Some philosophers have argued that value or goodness is a simple non-natural property of things (see Moore 1903 for the classical statement of this position), and others have argued that value or goodness is not a simple property of things but one that supervenes on the natural properties to which we appeal to explain a thing's goodness (this view is defended by, among others, contemporary moral realists; see McDowell, J. (1985). Values and Secondary Qualities. Morality and Objectivity. T. Honderich, Routledge and Kegan Paul**:** 110-29., Sturgeon, N. (1985). Moral Explanations. Morality, Reason, and Truth. D. C. a. D. Zimmerman, Rowman and Allenheld**:** 49-78; Sayre-McCord, G. (1988). The Many Moral Realisms. Essays on Moral Realism. G. Sayre-McCord. Ithaca, Cornell University Press**:** 1-26;, Brink, D. O. (1989). Moral Realism and the Foundation of Ethics. Cambridge, Cambridge University Press. .

[8] A large literature exists on the use of economic valuation methods to estimate the value of changes in environmental quality. For a comprehensive description of these methods, see Freeman (2003).

[9] In order to address these concerns, one or both of two steps can be taken in measuring public opinion. First, the EPA can conduct surveys that educate survey respondents about a problem or policy before asking for their preferences. Alternatively, in addition to preferences, the EPA can routinely measure the degree of knowledge and thought that people bring to a survey. This latter approach could be used to determine whether the preferences of members of the public who are knowledgeable and thoughtful are substantially different from those who are not. If this is the case, the EPA could conduct statistical simulations to predict what preferences people would express if everyone were fully informed and deeply thoughtful.

[10] In particular, the circular states that "'Opportunity cost' is the appropriate concept for valuing both benefits and costs. The principle of 'willingness-to-pay' (WTP) captures the notion of opportunity by measuring what individuals are willing to forgo to enjoy a particular benefit. In general, economists tend to view WTP as the most appropriate measure of opportunity cost, but an individual's 'willingness-to-accept' (WTA) compensation for not receiving the improvement can also provide a valid measure of opportunity cost" (OMB, p. 18).

[11] These interviews were conducted by one committee member, Dr. James Boyd, in conjunction with the Designated Federal Officer, Dr. Angela Nugent, over the period September 22, 2004 through November 23, 2005. In seven sets of interviews, Dr. Boyd spoke with staff from the Office of Policy, Economics and Innovation, Office of Water, Office of Air and Radiation, and the Office of Solid Waste and Emergency Response.

[12] NCEE is typically brought in by the program offices to both help design and review RIAs. NCEE can be thought to provide a centralized "screening" function for rules and analysis before they go to OMB. NCEE is actively involved in discussions with OMB as rules and supporting analysis are developed and advanced.

[13] In addition, the Circular states (p.27 – check page number) "If monetization is impossible, explain why and present all available quantitative information" and "If you are not able to quantify the effects, you should present any relevant quantitative information along with a description of the unquantified effects, such as ecological gains, improvements in quality of life, and aesthetic beauty" (add page number).

[14] The Committee reviewed and critically evaluated the CAFO Environmental and Economic Benefits Analysis at its June 15, 2004 meeting. As stated in the Background Document for SAB Committee on Valuing the Protection of Ecological Systems and Services for its Session on June 15, 2004, the purpose of this exercise was "to provide a vehicle to help the Committee identify approaches, methods, and data for characterizing the full suite of ecological 'values' affected by key types of Agency actions and appropriate assumptions regarding those approaches, methods, and data for these types of decisions." The Committee based its review on EPA's final benefits report (EPA 2002) and a briefing provided by the EPA Office of Water staff. During the June meeting, members of the Committee divided into two workgroups. The workgroups each worked independently and reported their findings to the combined Committee. The leaders of the two working groups then prepared a consolidated summary of comments from the two workgroups.

[15] In December 2000, EPA proposed a new CAFO rule under the federal Clean Water Act to replace 25-year-old technology requirements and permit regulations (66FR 2959). EPA published its final rule in December 2003 (68 FR 7176). The new CAFO regulations, which cover over 15,000 large CAFO operations, reduce manure and wastewater pollutants from feedlots and land applications of manure and remove exemptions for stormwater-only discharges.

[16] Prior to publishing the draft CAFO rule in December 2000, EPA spent two years preparing an initial assessment of the costs and benefits of the major options. After releasing the draft rule, EPA spent another year collecting data, taking public comments, and preparing assessments of new options. EPA published its final assessment in 2003. An intra-agency team at EPA, including economists and environmental scientists in the Office of Water, Office of Air and Radiation, Office of Policy Economics and Innovation, and Office of Research and Development, worked on the benefit assessment. EPA also worked with the U.S. Department of Agriculture in developing the assessment. Dr. Christopher Miller of EPA's Office of Water estimated that EPA spent approximately $1 million in overall contract support to develop the benefit assessment. EPA spent approximately $250,000-$300,000 on water quality modeling as part of the assessment.

[17] The potential "use" benefits included in-stream uses (commercial fisheries, navigation, recreation, subsistence, and human health risk), near-stream uses (non-contact recreation, such as camping, and nonconsumptive, such as wildlife viewing), off-stream consumptive uses (drinking water, agricultural/irrigation uses, and industrial/commercial uses), aesthetic value (for people residing, working, or traveling near water), and the option value of future

services. The potential "non-use" values included ecological values (reduced mortality/morbidity of certain species, improved reproductive success, increased diversity, and improved habitat/sustainability), bequest values, and existence values.

[18] These benefits were recreational use and non-use of affected waterways, protection of drinking water wells, protection of animal water supplies, avoidance of public water treatment, improved shellfish harvest, improved recreational fishing in estuaries, and reduced fish kills.

[19] These include reduced eutrophication of estuaries; reduced pathogen contamination of drinking water supplies; reduced human and ecological risks from hormones, antibiotics, metals, and salts; improved soil properties from reduced over-application of manure; and "other benefits".

[20] EPA apparently conducted no new economic valuation studies (although a limited amount of new ecological research was conducted) and did not consider the possible benefits of developing new information where important benefits could not be valued in monetary terms based on existing data.

[21] For example, while the report notes the potential effects of discharging hormones and other pharmaceuticals commonly used in CAFOs into drinking water sources and aquatic ecosystems, the nature and possible ecological significance of these effects is not adequately developed or presented. Similarly, the report does not adequately address the well-known consequences of discharging Trihalomethane precursors into drinking-water sources.

[22] In the case of this CAFO rule, 97% of the monetized benefits arise from recreation (boating, swimming and fishing) and from private well owners' willingness to pay for water quality, estimated using contingent valuation or travel cost methods.

[23] EPA used estimates based on a variety of public surveys in its benefit transfer efforts, including: a national survey (1983) that determined individuals' willingness to pay for changes in surface water quality relating to water-based recreational activities (Section 4 of the CAFO Report); a series of surveys (1992, 1995, 1997) of willingness to pay for reduced/avoided nitrate (or unspecified) contamination of drinking water supplies (Section 7); and several studies (1988, 1995) of recreational fishers' values (travel cost, random utility model) for improved/protected fishing success related to nitrate pollution levels in a North Carolina estuary (Section 9).

[24] Although EPA later prepared more detailed conceptual models of the CAFO rule's impact on various ecological systems and services, EPA did not prepare these models until after the Agency finished its analysis.

[25] Contamination of estuaries, for example, might negatively affect fisheries in the estuary (a primary effect) but might have an even greater impact on offshore fisheries that have their nurseries in the estuary (a secondary effect).

[26] The goal of EPA's analysis was a national level assessment of the effects of the CAFO rule. This involved the effects of approximately 15,000 individual facilities, each

contributing pollutants across local watersheds into local and regional aquatic ecosystems. A few intensive case studies were mentioned in the report and used to calibrate the national scale models (e.g., NWPCAM, GLEAMS), but there was no indication that these more intensive data sets were strategically selected or used systematically for formal sensitivity tests or validations of the national-scale model results.

[27] This could include either a robust public involvement process following Administrative Procedures Act requirements (e.g., FR publication), or some other public involvement process [see EPA's public involvement policy, U.S. Environmental Protection Agency Office of Policy, E. a. I. (2003). Public Involvement Policy of the U.S. Environmental Protection Agency. EPA 233-B-03-002.; the SAB report on science and stakeholder involvement U.S. Environmental Protection Agency Science Advisory Board (2001). Improved Science-Based Environmental Stakeholder Processes: An EPA Science Advisory Board Commentary. EPA-SAB-EC-COM-001-006.

[28] Models may be valuable in many of the steps of assessing ecological value including: estimating stress loading; estimating the exposure pattern of stress – especially spatial and temporal implication; identifying Ecological element(s) receiving exposure; estimating exposure - response function of ecological elements; estimating the reduction or prevention of increased stress from agency action; estimating the response of service production or function to change in stress; valuating the ecological service associated with that change in production; linking to economic or socio-political for further valuation in dollars or other metrics

[29] Note that these essential ecosystem characteristics are very similar to the seven ecological indicators in EPA's report on assessing ecological systems: landscape condition, biotic condition, chemical and physical characteristics, ecological processes, hydrology and geomorphology and natural disturbance regimes, Young and Sanzome, 2002

[30] This analysis evaluated the benefits and costs of amendments to the Clean Air Act passed by Congress in 1990. Its effort to evaluate the ecological benefits of these amendments raises many of the same issues that arise in evaluating the benefits of national rules. In the prospective analyses the sequence of increasingly stringent rules called for under the 1990 Clean Air Act Amendments are compared with a situation where the rules were held constant at their 1990 levels (e.g. with the regulatory regime prior to the amendments).

[31] A number of the gasses emitted from CAFOs have adverse air quality impacts that are interrelated with the water quality impacts.

[32] The pollutants that result from CAFOs have environmental effects that are local, regional and global. For example, in terms of emitted gases, methane and N20 are major greenhouse gasses of global concern; ammonia and NOx have important regional impacts on air quality and nitrogen deposition; and odor and suspended particulate matter have important local or on-site impacts (NRC,2003)

[33] In the case of air, nitrous oxide has a lifetime of 100 years in the planetary boundary layer, whereas hydrogen sulfide has a lifetime of only about a day. These spatial and temporal dimensions of dispersion and lifetime of effects also apply to many of the water pollutants

although the spatial dimensions do not extend to the global.

[34] CAFOs are not uniformly distributed in the country or even within a state. For various reasons they often are clustered. Each of these concentration areas has unique climatic, soil and topographic features that influence waste dispersion. Further, manure type, in addition to soil characteristics, has a differential impact on soil microbial populations and hence on decomposition rates (Larkin 2006).

[35] The animal feed used at CAFOs no longer comes from local surroundings but may be produced in areas remote from the sink facilities, including foreign sources. The production of these grain feeds results in non-point pollution in the production regions. Further, fish meal is an important feed supplement for pigs and chickens with the fish generally being harvested from coastal and marine ecosystems, often from places far distant from the United States, with consequences for local food chains.

[36] A syndrome has been identified that involves: increased biomass of phytoplankton, shifts in phytoplankton to bloom-forming species that may be toxic, in marine environments, increases of gelatinous zooplankton, increases in biomass of benthic and epiphytic algae, changes in macrophytic species composition, decreases in water transparency, oxygen depletion, increased incidence of fish kills, and loss of desirable fish species (Carpenter et al., 1998). There are a number of important features of this syndrome. It is easily recognized, it is reversible, and there are some features that show up early and hence provide indicators of ecosystem disruption and early opportunities for mitigation. Clean water and recreational opportunities have been extensively treated in valuation projects. The impacts on the biological nature of a system may not be readily appreciated or valued by the public, but it certainly provides an indicator that the things they do value are in trouble. The power of public involvement in understanding, valuing and responding to eutrophication is shown by the classic example from Lake Washington (Smith, 1998). The understanding part took considerable efforts in educating the public by those few scientists who understood what was happening.

[37] See Phaneuf and Smith(2005) for a review of the literature and Phaneuf (2002, Phaneuf, Palmquist and Smith (2006), (Egan (2004), von Haefen (1999), and Egan and Herriges (2006) for examples of applications involving freshwater recreation sites in different regions.
[38] Consumer surplus measures the excess of the sum of the marginal values over the expenditures that must be made to obtain the good at a fixed price. Thus, consumer surplus sums up the differences between the maximum a consumer would be willing to pay for a good minus the amount actually paid (price) for each unit consumed. Similarly, producer surplus measures the excess of receipts for the good over the sum of the marginal costs to provide each unit. Producer surplus is then a comparable concept. It aggregates the difference between what producers are willing to sell a product for (supply) and what they actually receive (price) for each unit they provide. Adding together changes in consumer surplus and producer surplus generates the change in total economic benefit.

[39] The last component of these costs, the cost of time on site per visit, is difficult to include because it is reasonable to assume it is jointly determined with decisions about the location to visit and the number of trips to take in a season . It is also related to measures of the amount of the site's services that are consumed. Most studies acknowledge these costs as an issue but

don’t include them in the analysis as a result of these difficulties. As a rule the time on site per trip is assumed to be held constant.

[40] Correlations can range from 1 (meaning a perfect match between the variables) to 0 (meaning a relation between the variables no better than chance) to -1 (meaning a perfect inverse relation between the variables).

[41] Another relevant literature travels under the heading of “pretest sensitization” effects (see Bracht & Glass, 1968; Welch & Walberg, 1972; Willson & Putnam, 1982). Experimental researchers conducting within-subjects designs have considered the possibility that merely measuring their dependent variable in a pretest might influence their dependent variable. In particular, there has been an interest in the possibility that measuring an attitude or belief may create inertia in it, making this construct more resistant to change later. Consistent with this notion, Nosanchuk and Marchak (1969) found that expressing opinions on issues in a questionnaire reduced the amount of attitude change that participants manifested later in response to persuasive paragraphs. And Jagodzinski, Kuhnel, and Schmidt (1987) found that interviewing people on a topic made their attitudes on it more stable subsequently.