UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
WASHINGTON D.C. 20460

OFFICE OF THE ADMINISTRATOR
EPA SCIENCE ADVISORY BOARD

The Honorable Lisa P. Jackson
Administrator
U.S. Environmental Protection Agency
1200 Pennsylvania Ave., NW
Washington, D.C. 20460

Subject: Review of *Particulate Matter Urban-Focused Visibility Assessment (External Review Draft, September 2009)*

Dear Administrator Jackson:

The Clean Air Scientific Advisory Committee (CASAC) Particulate Matter (PM) National Ambient Air Quality Standards (NAAQS) Review Panel met on October 5 - 6, 2009 to review the *Particulate Matter Urban-Focused Visibility Assessment* (External Review Draft, September 2009). In this letter, CASAC offers general comments on the *Visibility Assessment*, followed by consensus responses to the Agency's charge questions. Comments from individual panelists are also enclosed.

The *Urban-Focused Visibility Assessment* provides a sound basis for EPA to consider a new secondary PM standard based on visibility as a welfare endpoint. CASAC was impressed by the comprehensive approach, quality of analysis, and clarity of presentation evident in the draft assessment. While the *Visibility Assessment* focused only on selected urban areas, a new secondary PM standard would apply to all urban and rural areas of the country, and provide an effective complement to the protection currently provided in Class 1 Federal National Parks and Wilderness Areas by the 1999 Regional Haze Rule. CASAC strongly supports the introduction of a new PM light extinction indicator for a secondary standard to protect against adverse effects on visibility. We also concur that an hourly averaging time most appropriately represents the nearly instantaneous nature of human perception of impaired visibility. The 20 to 30 deciview (DV) range of light extinction levels evaluated in the assessment is a good representation of the levels of visibility impairment considered to be unacceptable based on the urban visibility preference studies currently available. CASAC also considers this to be a reasonable range for the Administrator to consider in selecting a new secondary PM NAAQS. We encourage the Agency to conduct additional visibility preference studies prior to the next PM NAAQS review to support refinements of this standard in future PM NAAQS review cycles.

To evaluate the PM light extinction indicator, EPA chose the 90th and 95th percentiles of the daily maximum daylight hours as the form for this indicator. These percentiles were used to evaluate relationships among various levels for primary $PM_{2.5}$ mass standards and new secondary PM light extinction standards under consideration as "candidate protection levels." However, neither of these percentiles can be supported as the single best choice of the form of the standard, and CASAC recommends that other percentiles be considered, up to and including the 98th percentile used for the current 24-hour primary and secondary standards. CASAC also recommends lowering the relative humidity (RH) limit from 95% to 90% and using it as a screen (i.e., hours above it should be discarded) rather than a cap, to more clearly exclude weather events like fog or precipitation and to minimize effects of measurement error and spatial variability.

Visibility varies between the more humid and often more polluted Eastern U.S. and the Western U.S.; however, the *Visibility Assessment* demonstrates that current visibility conditions in all the urban study areas would sometimes fail to meet even the most lenient candidate protection level under consideration (a PM extinction level of 30 DV). In general, the more humid Eastern urban areas would have more days that fail to meet the standard. In fact, Eastern urban areas would exceed a 30 DV threshold on an average of 20% or more of the days even after assuming full compliance with the most stringent primary PM standard under consideration ($12\mu g/m^3$ and $25\mu g/m^3$). While Eastern states (with the exception of southern California) tend to have poorer visibility, they are also more populated, and thus visibility impairment affects a larger number of people. In the Western U.S., small increases in pollutant concentrations can markedly degrade visibility. In addition, the *Visibility Assessment* suggests that there may be a preference for better visibility levels by residents in some Western US urban areas.

To ensure progress in both Western and Eastern states, CASAC suggests that EPA consider a "progress based" metric rather than a "threshold based" standard. A progress-based secondary PM NAAQS would incorporate elements of the traditional standard setting process, retaining the indicator, averaging time, form and level. Compliance with the standard would then be based on maintaining a minimal percentage reduction over time in the frequency of daylight hours or days exceeding the level. A standard requiring a uniform rate of progress over time would encourage a reasonable rate of improvement in the East, while still providing protection and incentives for improvement in the West. An ongoing, steady, progress-based approach would be more directly consistent with, and could be much more efficiently coordinated with, the Regional Haze Rule requirements for reasonable rates of progress in improving visibility in all (Eastern and Western) National Parks and Wilderness Areas with Class I status, regardless of the current baseline levels of impairment. Moreover, it would be consistent with the absence of fixed time requirements for attaining secondary standards, which are to be achieved "as expeditiously as practicable."

Another approach would be to set the standard at a relatively stringent level and percentile, such that it would be protective in the West, with the clearly stated intention to specify guidance on future rates of progress that will be considered reasonable and which would be intentionally less rapid for urban areas than for the protected National Parks and Wilderness areas covered by the Regional Haze Rule. For the upcoming Policy Assessment, we ask that

EPA consider methods for assessing progress and achieving the various secondary PM NAAQS options, in addition to the form, level, averaging time, and indicator options.

Finally, EPA should carefully consider the new monitoring program that will be required to implement a visibility standard. While there are a variety of commercially available instruments capable of measuring hourly light extinction and/or its separate components of light scattering and absorption, such measurements are currently not being made in most urban areas or in large-scale network operations. For this reason, CASAC strongly encourages the Agency to move quickly to establish a small, urban visibility pilot monitoring network to test, evaluate and refine current and evolving light extinction measurement methods. The Agency is also encouraged to seek additional advice from the CASAC Ambient Air Monitoring and Methods Subcommittee (AAMMS) on the most suitable methods for a national urban visibility monitoring network.

We thank the Agency for the opportunity to provide advice on the PM Urban-Focused *Visibility Assessment* and look forward to the review of EPA's Policy Assessment early next year.

Sincerely,

Enclosure A: CASAC Particulate Matter Review Panel Roster
Enclosure B: CASAC Responses to Charge Questions
Enclosure C: Individual Panelists' Responses to Charge Questions

Enclosure A

Clean Air Scientific Advisory Committee
Particulate Matter Review Panel

CHAIR

Dr. Jonathan M. Samet, Professor and Chair, Department of Preventive Medicine, University of Southern California, Los Angeles, CA

CASAC MEMBERS

Dr. Joseph Brain, Philip Drinker Professor of Environmental Physiology, Department of Environmental Health, Harvard School of Public Health, Harvard University, Boston, MA

Dr. Ellis B. Cowling, University Distinguished Professor At-Large Emeritus, Colleges of Natural Resources and Agriculture and Life Sciences, North Carolina State University, Raleigh, NC

Dr. James Crapo, Professor of Medicine, Department of Medicine, National Jewish Medical and Research Center, Denver, CO

Dr. H. Christopher Frey, Professor, Department of Civil, Construction and Environmental Engineering, College of Engineering, North Carolina State University, Raleigh, NC

Dr. Donna Kenski, Data Analysis Director, Lake Michigan Air Directors Consortium, Rosemont, IL

Dr. Armistead (Ted) Russell, Professor, Department of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA

CONSULTANTS

Dr. Lowell Ashbaugh, Associate Research Ecologist, Crocker Nuclear Lab, University of California, Davis, Davis, CA

Prof. Ed Avol, Professor, Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA

Dr. Wayne Cascio, Professor, Medicine, Cardiology, Brody School of Medicine at East Carolina University, Greenville, NC

Dr. David Grantz, Director, Botany and Plant Sciences and Air Pollution Research Center, Riverside Campus and Kearney Agricultural Center, University of California, Parlier, CA

Dr. Joseph Helble, Dean and Professor, Thayer School of Engineering, Dartmouth College, Hanover, NH

Dr. Rogene Henderson, Senior Scientist Emeritus, Lovelace Respiratory Research Institute, Albuquerque, NM

Dr. Philip Hopke, Bayard D. Clarkson Distinguished Professor, Department of Chemical Engineering, Clarkson University, Potsdam, NY

Dr. Morton Lippmann, Professor, Nelson Institute of Environmental Medicine, New York University School of Medicine, Tuxedo, NY

Dr. Helen Suh MacIntosh, Associate Professor, Environmental Health, School of Public Health, Harvard University, Boston, MA

Dr. William Malm, Research Physicist, National Park Service Air Resources Division, Cooperative Institute for Research in the Atmosphere, Colorado State University, Fort Collins, CO

Mr. Charles Thomas (Tom) Moore, Jr., Air Quality Program Manager, Western Governors' Association, Cooperative Institute for Research in the Atmosphere, Colorado State University, Fort Collins, CO

Dr. Robert F. Phalen, Professor, Department of Community & Environmental Medicine; Director, Air Pollution Health Effects Laboratory; Professor of Occupational & Environmental Health, Center for Occupation & Environment Health, College of Medicine, University of California Irvine, Irvine, CA

Dr. Kent Pinkerton, Professor, Regents of the University of California, Center for Health and the Environment, University of California, Davis, CA

Mr. Richard L. Poirot, Environmental Analyst, Air Pollution Control Division, Department of Environmental Conservation, Vermont Agency of Natural Resources, Waterbury, VT

Dr. Frank Speizer, Edward Kass Professor of Medicine, Channing Laboratory, Harvard Medical School, Boston, MA

Dr. Sverre Vedal, Professor, Department of Environmental and Occupational Health Sciences, School of Public Health and Community Medicine, University of Washington, Seattle, WA

SCIENCE ADVISORY BOARD STAFF

Dr. Holly Stallworth, Designated Federal Officer, EPA Science Advisory Board Staff Office

NOTICE

This report has been written as part of the activities of the EPA's Clean Air Scientific Advisory Committee (CASAC), a federal advisory committee independently chartered to provide extramural scientific information and advice to the Administrator and other officials of the EPA. CASAC provides balanced, expert assessment of scientific matters related to issues and problems facing the Agency. This report has not been reviewed for approval by the Agency and, hence, the contents of this report do not necessarily represent the views and policies of the EPA, nor of other agencies within the Executive Branch of the federal government. In addition, any mention of trade names or commercial products does not constitute a recommendation for use. CASAC reports are posted on the EPA Web site at: http://www.epa.gov/casac.

Enclosure B

CASAC Responses to Charge Questions on the *Urban-Focused Visibility Assessment* (External Review Draft, September 2009)

1) After careful consideration of the evidence provided in the second draft ISA, and in particular the significant body of work that has been conducted by the Regional Planning Organizations under the Regional Haze Rule, (i.e. information on urban and rural PM concentrations and compositions), we have decided to continue to focus this assessment on the PM induced visibility impairment that is occurring in urban areas. What is the Panel's view on this approach? Is the rationale supporting the selected approach clear and appropriate?

CASAC found the rationale for pursuing a secondary PM NAAQS based on daytime urban visibility to be well justified, based on previous work associated with the Regional Haze Rule, on the existing ability to model this parameter based on available PM mass measurements, and on the concentration of impacted populations in these areas. There was some concern that visibility should not be assumed to represent all of welfare effects of PM. even though it is analytically more accessible at this time.

Because urban areas represent only a small fraction of the land mass of the U.S., a meaningful secondary standard, complementary to the Regional Haze Rule, should extend to rural non-Class I areas.

Some aspects of the assessment require more rigorous presentation. Uncertainties surrounding nitrate mass, including volatilization, presence in coarse and fine modes, and degree of neutralization, for example, require further evaluation and explanation in the document. The specific quantitative application of measured relative humidity, the method of determining daylight hours, and the actual method used to calculate policy relevant background could be more fully presented.

2) After further considering the nature of urban versus more remote area PM, and in light of discussions with CASAC at the April 2, 2009 meeting, we have decided not to develop an urban optimized algorithm at this time, but instead to rely on the original IMPROVE algorithm to relate urban PM to local haze (PM light extinction). Is the Panel generally supportive of this approach? Is the rationale supporting this decision appropriate and clearly presented?

The decision to use the IMPROVE algorithm to relate urban PM to local haze is suitable . Given the underlying uncertainties such as R_{oc} factors for urban environments, a lack of detailed PM composition information and other biases in estimating species concentrations in urban areas, and the lack of optical measurements in urban areas, the approach taken and presented in Chapter 1 is reasonable.

It is recommended that the original IMPROVE equation be modified to include sea salt, and consideration should be given as to what the appropriate R_{oc} factor might be.

The rationale supporting this approach is appropriate and clearly articulated in the document.

3) In a change from the planned approach presented in the Scope and Methods Plan, we have decided to conduct a reanalysis of the urban visibility preference studies available at the time of the 2006 PM NAAQS review, rather than conducting new public preference studies since it is highly unlikely that the results of new studies could be completed in time to inform this review. This reanalysis was designed to explore the similarities and differences (comparability) between the current studies and to assess what information could be drawn from these results to inform the selection of visual air quality (VAQ) candidate protection levels (CPLs) to be used in subsequent impact assessments. This reanalysis also includes a recent study by Smith and Howell (2009) for Washington, D.C. which was presented to the CASAC during the public comment phase of the April 2, 2009 meeting and later provided to EPA staff. Does the Panel agree that the information provided by this reanalysis is useful to inform the selection of CPL? Does the Panel agree that inclusion of the Smith and Howell (2009) is appropriate in both the ISA and Visibility Assessment? To what extent does the Panel consider that the reanalysis of the urban visibility preference studies is clearly and appropriately characterized?

For the purposes of evaluating a secondary PM NAAQS protective of non-Class I area visual air quality in this PM NAAQS review cycle, we support the reanalysis effort in lieu of conducting additional preference studies, given the time available. We support the inclusion of the Smith and Howell data. As noted in the answer to Charge Question 8, to better understand and bound uncertainty, a logistic distribution could be fitted to the data and percentile values extracted from the curve fit, along with associated confidence intervals. This analysis would provide an understanding as to whether the values derived from various studies are indeed statistically different from each other.

While the Agency is encouraged to allocate sufficient funds and resources to support the planned new urban visibility preference studies, there would not have been time to complete the new studies quickly enough to support the current assessment, and the results of those new studies will be more useful if they are carefully planned and executed, and not rushed to meet an accelerated NAAQS review schedule. The reanalysis shows a relatively strong degree of convergence in the identification of unacceptable levels of visual air quality across the different study areas, and provides a sufficient basis for use of the proposed CPLs in the current risk assessment, and for considering alternative levels of secondary PM standards based on PM light extinction.

4) We have chosen to use the range that represents the 50th acceptability criteria across the four cites studied (i.e., the VAQ level that best divides the photographs shown in into two groups: those with a VAQ rated as acceptable by the majority of the participants, and those rated not acceptable by the majority of participants) as CPLs to characterize the nature of the impact on urban VAQ associated with current PM levels. Please comment on the clarity and appropriateness of the rationale supporting this decision. Does the Panel have suggestions for alternative ranges to consider?

CASAC agreed that the presentation is clear, and that the 50th percentile for the acceptability criteria is logical, given the noted similarities in methodologies employed in the 4 study areas.

The reanalysis could also be done at the 90th percentile for acceptable visibility, as has been done in various perception threshold studies. In terms of choosing a specific percentile from the preference studies, we note that there may not be a "preferred" one, but in assessing preference studies to propose a PM secondary NAAQS, the 50th percentile is sufficient, as it is the basis for existing visibility indexes used in the Denver/Colorado Front Range and Phoenix metropolitan areas. It may miss shifts in other parts of the distributions for different parts of the country as emissions controls are applied. It is likely that these "acceptable" and "unacceptable" VAQ levels divided at the 50th percentile will be further refined in the future, as results of additional studies become available. For use in the current NAAQS review, the identified range of VAQ levels provides a sound and defensible basis for establishment of a separate PM secondary standard based on achieving and maintaining acceptable levels of VAQ in urban areas.

The Phoenix study with the large sample size confirms the results from the more limited sample sizes in the other studies. Each study incorporated various visibility ranges to arrive at the unique 50th percentile values. Future visibility preference studies of large unbiased sample sizes would assist in further confirming and refining these results for future NAAQS review cycles. These studies should include variation in urban and rural locations, contributing sources, meteorology, climate, population size, aerosol characteristics, and monitoring/viewing site representativeness across the U.S.

It should also be noted that for all the visibility preference studies reviewed here, respondents were specifically directed to exclude consideration of potential health effects in making their judgments. In reality, human perception is complexly related to feelings of psychological wellbeing and/or emotional stress, which in turn can predispose or enhance the direct health effects from exposures to visibility-impairing pollutants. More work is needed to better understand these complex visibility/health relationships, as their artificial exclusion from past preference studies has likely resulted in apparent tolerance for poorer levels of VAQ than would otherwise be considered acceptable.

5) A new indicator relating ambient PM to urban VAQ (i.e., PM light extinction) has been evaluated to improve our characterization of the relationship between ambient PM and visibility impairment. Is the Panel generally supportive of this approach? To what extent have we provided an adequate justification for the new indicator used?

The use of visual air quality as an indicator of urban VAQ is reasonable, and is an approach CASAC strongly supports. Adequate justification for the use of a direct measurement is provided. Light extinction is an easily measured quantity that can be related in a reasonable way to observed visibility so that it provides an objective measure of VAQ. CASAC recommends that an RH limit be used as a screen (hours above it should be discarded) rather than a cap (where RH levels for hours above the cap are set equal to the cap), and also recommends that the RH screen be lowered to 90%. This lower RH limit would unavoidably eliminate consideration of some of the worst visibility conditions, but would provide greater confidence that degraded visibility was a result of pollution and not weather.

6) **An averaging time of one hour as a practical minimum time period was used in this assessment in recognition that, while the visibility impacts are nearly instantaneous, the urban VAQ does not generally change significantly from minute to minute, but does vary from hour to hour. To what extent does the Panel support this approach? Does**

the Panel consider the rationale supporting this approach to be clearly and appropriately presented?

One hour averaging is a reasonable compromise, and the presentation in the document is credible. It is likely that in most circumstances, such an averaging period would be reasonable and matches the time frame for measurement of other pollutants. As indicated in the response to question # 7 below, a focus on the single worst hour of the day is not necessarily the best way to employ an hourly averaging time as the basis for a regulatory metric.

7) We have chosen to use the 90th and 95th percentile forms in our assessment of alternative secondary (welfare-based) standards. Please comment on the use of these alternative forms.

Assessing the choice of a 90th, 95th, or 98th percentile form cannot really be done independently of the choice of Candidate Protection Levels (CPLs). We have taken the liberty of exploring the implications of the three candidate CPLs of 20 dv (74 Mm^{-1}), 25 dv (122 Mm^{-1}), and 30 dv (201 Mm^{-1}) in combination with the 90th, 95th, and 98th extinction percentiles, using extinction data that have been reconstructed from IMPROVE-type measurements at Atlanta, Georgia (see chart). The data are 24-hr averages, so the analysis is not directly applicable to setting a 1-hr standard but does serve to highlight a few key points. We have used a 95% RH cutoff with all RH values above 95% set equal to 95%.

Table 1: Percentile Form of PM Secondary NAAQS and "Number of Days" Affected by Candidate Protection Levels (from example 24-hour Urban Light Extinction Data chart at end)

	Candidate Protection Levels[1]		
	20 deciviews (74 Mm^{-1})	**25 deciviews (122 Mm^{-1})**	**30 deciviews (201 Mm^{-1})**
# of days from example dataset that exceed the 20, 25, and 30 deciview CPL (observations adjusted to 365 days)	**299**	**172**	**51**
# of days to be moved from Unacceptable to Acceptable at 90th Percentile (observations adjusted to 365 days)	**263**	**136**	**15**
# of days to be moved from Unacceptable to Acceptable at 95th Percentile (observations adjusted to 365 days)	**281**	**154**	**33**
# of days to be moved from Unacceptable to Acceptable at 98th Percentile (observations adjusted to 365 days)	**292**	**165**	**44**

1 - From reanalysis of 4 visibility preference studies, applied to 15 cities' fine mass data, converted to reconstructed light extinction

Table 1 summarizes the results of the analysis. First, choosing the 90th, 95th, or 98th percentile corresponds to allowing 36, 18, and 7 days to exceed the target (standard) level. Currently, the 24-hr extinction values exceed the 20 dv, 25 dv, and 30 dv on 299, 172, and 51 days of the year, based on the example. Keep in mind that these are for 24-hr extinction values, and undoubtedly the frequency of exceedances will be greater for 1-hr data. Notice that the number of days

required to move from the unacceptable to acceptable for the 30-dv standard and for the 98th, 95th, and 90th percentiles is 44, 33, and 15 days, while for the 20-dv standard it is 292, 281, and 263 days. The 25-dv CPL is between these two extremes. Clearly, the choice of the level of the standard has a more significant effect on implied emission reductions than the choice of 98th, 95th, or 90th percentile. Secondly, if the b_{ext} values are lognormally distributed, the 98th percentile will be more restrictive than the 95th, which in turn is more restrictive than the 90th percentile. If the distribution is not lognormal, it is possible that the 98th percentile would target a single source or source area, while the 90th percentile choice would be responsive to more regional transport. As stated above, the analysis was carried out on a 24-hr dataset and with a 95% RH round-down assumption. The same analysis using 1-hr data will push more extinction values above a chosen exceedance value, while choosing a 90% RH cutoff will reduce the extinction values and tend to reduce the number of exceedances.

The assessment presented in the Urban Visibility Assessment document assumed a distribution of 1 hr maximum daylight hours. Another alternative would be to assume all daylight hours from which to pick the underlying distribution without concern for daily maximums. What are the potential effects of these choices on emission control strategies - will they be any different? Again, choosing the distribution from all daylight hours versus the maximum from each day may target single-source areas, while a control strategy responsive to improving the maximum extinction on each day may very well target sources having persistent regional contributions. Because the underlying distributions of extinction as well as the aerosol compositional structure of the distributions will vary significantly from urban area to urban area, it might be preferable to choose a rate-of-progress standard as opposed to an absolute standard. This would allow picking a relatively low dv level, such as 20 dv, that is not only protective of the eastern but also of the western cities. The standard could then be structured so as to require a 10% reduction per planning period in exceedances at all urban areas. This approach could require bringing 10 days below the exceedance value in an eastern city, while only requiring 1 day for a western urban area. All cities would have to make some ongoing progress over time toward improved visibility.

We suggest exploring a number of reasonable alternative strategies in order to better characterize the outcomes with regard to control and enhancement of visibility over time.

8) To what extent does the Panel support the graphical displays presented in this chapter? As currently presented, do these figures clearly summarize the assessment results? We have combined data from multiple studies for two locations - British Columbia and Washington, DC - (Abt, 2001; Smith and Howell, 2009 test 1) as presented in figure 2-14. Does the Panel agree with developing a composite dataset for each of these two urban areas?

We note that all these studies are somewhat different in methodologies and results, while sharing several noted similarities. While the emissions source mixes, ambient aerosol loadings and compositions, climate, and meteorology are quite different across these 4 study areas that produce the 20 to 30 deciview preference studies' results, we conclude that this range of values is sufficiently tested and reasonable, for the purposes of selecting the Candidate Protection Levels defined at the end of Chapter 2.

The graphic displays (such as those in Figures 2-2, 2-3, 2-5, 2-7, and 2-9 through 2-14) which compare visibility levels in deciviews with percentages of respondents rating the VAQs acceptable serve well in showing results from the different studies in common units, showing the range or distributions among different respondents and across the different study areas, and showing the convergence of results for acceptable/unacceptable visibility in the range of 20 to 30 dv. We would also like to see the absolute change in 1/Mm also shown on X axis below the deciview scale. The use of composite data sets based on combined study results for the British Columbia studies and for the Washington DC studies is reasonable, and seems appropriate for combining results from these study areas with those from the Phoenix and Denver studies.

A composite of findings from the urban areas where studies were conducted and indices selected (Denver, Phoenix) with those having only study results (Lower Fraser Valley, Washington, DC) does not offer a direct comparison. The indices involve further analysis and technical/operational considerations beyond the visibility preference studies' results. Combining of the various preference studies is acceptable, as long as the measure of the studies' results is the value for the 50% acceptability criteria and not the ultimate index values selected. Further, a logistic distribution could be fitted to the data and percentile values extracted from the curve fit, along with associated confidence intervals. This analysis would provide insight as to whether the values derived from the various studies differ statistically.

9) Despite significant differences in study characteristics (e.g., size, location), to what extent does the Panel support combining and comparing the results from the four cities, as shown in Figure 2-14? What is the Panel's view on the clarity and adequacy of the descriptions of the uncertainties and limitations associated with such a combined assessment and the conclusions that can be drawn from the assessments? Please provide comments on additional insights, uncertainties, or caveats that should be considered.

Combining and comparing the results from the four different study areas is an appropriate approach that takes maximum advantage of all the currently available data. To a certain extent, the relative convergence of results from the different study approaches and study areas strongly supports the use of these data collectively for use in identifying a relatively narrow range of acceptable VAQs across many locations and respondents. Although a formal uncertainty analysis was not conducted, the descriptive information on uncertainties and limitations of this approach conveys that information clearly. The analysis using a logistic distribution mentioned under #8, above, would assist in narrowing and bounding uncertainty.

It should also be noted that the uncertainties associated with this approach for selecting a level or range of acceptable VAQ are likely to be offset by, and/or interactive with, the uncertainties associated with other aspects of a secondary PM standard - such as the averaging time, representing a day by its worst 1 (or more) VAQ hour(s), and the frequencies (percentiles) for which unacceptable VAQ conditions are acceptable. While the focus here has been on discerning a decision point or range of VAQ conditions considered acceptable versus unacceptable, it should also be recognized that there are welfare benefits associated with improvements across all parts of the visibility spectrum, including, for example, a shift of days with "good" to "excellent" VAQ. As a practical matter, efforts made to improve visibility

conditions in any part of the distribution are likely to shift the entire distribution, with benefits that will be understated by a single 50% acceptability approach. The studies' results inform the development of a consistent national secondary PM standard that may begin to protect VAQ using a light extinction metric (the quantitative value assessed in the preference studies).

10) We have used the combined results presented in this chapter to develop a range of CPLs that are used in subsequent steps of the assessment. To what extent does the Panel support the range of CPLs used and the justification provided for selecting this range? Does the Panel recommend consideration of any alternative approaches or criteria for selecting CPLs?

CASAC supports using combined results from existing urban visibility studies as a basis for selecting candidate protection levels (CPLs), which span the fairly narrow (20 to 30 dv) range of VAQ levels meeting the 50% acceptability criteria across the available study areas. As mentioned, the analysis could be done in a more elegant and statistically defensible way by fitting a logistic distribution to the data. Percentile values could then be extracted from the curve fit, along with associated confidence intervals. This analysis would help to show whether the values derived from various studies are indeed statistically different from each other, and percentiles other than the median could also be considered along with other elements of the proposed standard. It is probable that results from planned future visibility preference studies from other urban areas employing improved survey techniques may help narrow or otherwise refine the indicated 20 to 30 dv range of the CPLs proposed here. However, the currently proposed range is sufficiently well justified for use in the current standard setting process, in combination with other appropriate selections of indicator, averaging time and form.

11) Overall, we consider this assessment useful for providing information for the design of future urban visibility preference studies. Does the Panel support this conclusion and does the Panel have any recommendations for changes that could be made in the discussions of this information to enhance its usefulness for this purpose?

CASAC concurs that this assessment provides a useful basis for identifying designs for, and priority information needs from, future urban visibility studies. Such studies should be designed to provide information to help refine several of the elements of a PM light extinction standard in addition to identifying the level(s) of VAQ considered unacceptable across different urban locations and viewing conditions. For example, what frequencies of unacceptable visibility conditions are "acceptable" over time, and should these frequencies be applied to the single worst hour of a day or should all daylight hours be considered? In addition to a single "threshold" of unacceptable visibility conditions, what are preferences for shifts in other percentiles of the visibility frequency distribution – such as an increase in the frequency of very clear days? Future studies should also evaluate the potential role of aerosols on sky color and visibility of clouds, which are both integral to good visibility in an urban setting.

As always, there is a need to carefully balance the advocacy for future research needs with the confidence in using the best currently information to make sound decisions to better protect

human health, welfare and environment. In this case, currently available information is adequate to establish a basic secondary standard based on PM light extinction, even while future research is clearly needed to refine such a standard in the future. It should also be noted that the more detailed information on hourly light scattering and absorption that will result from implementing a standard with a PM extinction indicator, will itself provide an invaluable information resource to support future urban visibility preference studies and refine extinction equations for different mixes of urban aerosols. These data would also be available for near-real time reporting and public communication, will be useful for air quality forecasts and (in combination with other continuous measurements) will provide added information on highly time-resolved fine & coarse particle composition, including black carbon, of value for various health effects studies as well.

12) Are the goals articulated in the first paragraph of this chapter achieved in the remainder of the assessment? If not, does the Panel have suggestions for additional assessments that should be done?

13) Are the methods and approaches taken in these assessments, including those for monitor site selection, incomplete data adjustments, and the use of the CMAQ model to augment speciation data, appropriate and is the rationale for their selection clearly articulated?

14) Is the approach used to estimate PRB as described in chapter 3 and Appendix C appropriate?

Response to Questions 12 – 14: CASAC believes Chapter 3 provides a thorough assessment of the limitations of existing data and the uncertainties related to an assessment of visibility (light extinction). One addition may be appropriate – there are several monitoring networks that could serve as a validity check on the modeled results described in the assessment. Both SEARCH (Southeastern Aerosol Research Characterization Study) and the IMPROVE program (Interagency Monitoring of Protected Visual Environments) have samplers in urban areas, including some of the 15 cities examined here. The IMPROVE network routinely measures $PM_{10-2.5}$ mass, and SEARCH measures $PM_{10-2.5}$ using dichotomous virtual impactors for several species. Data from both networks might be useful for comparison with the results of the modeling method used here.

The approach taken to assess urban visibility is sound. Given the relative lack of measurements (except for the few sites that have IMPROVE samplers) the approach to obtaining estimates of hourly daytime visibility is a good one. The method is clearly articulated.

The approach used to estimate PRB is also sound. CASAC believes it may serve as a useful check on the modeled results to compare appropriate IMPROVE visibility measurements with the PRB results.

It appears that EPA generalized each “season” so that each location was defined to have the same number of daylight hours. It is not clear what benefit this provides, and the suggestion (footnote 23) that this may be eliminated in the final version of this assessment seems appropriate.

CASAC questions the use of a ratio of $PM_{10-2.5}$ to $PM_{2.5}$ to allocate hourly measurements. Both PM_{10} and $PM_{2.5}$ change rapidly, but PM_{10} does so especially. Whenever possible, it would be useful to use a ratio based on daily values to perform the allocation instead of the longer time average (monthly or annually) that was used here.

15) We consider the results generated by these analyses to be reasonable based on PM composition and relative humidity data. Does the Panel agree? Are there other tests of reasonableness that could be applied?

The results presented in this assessment are reasonable given the data that are available, the methods employed in the Assessment to estimate hourly light extinction, and the current state of understanding of the effects of PM composition and humidity on extinction. If time allows, CASAC recommends that EPA explore the use of independent data as consistency checks for the hourly estimates of species composition and/or light extinction. Examples of hourly PM composition data include the SEARCH sites in Atlanta and Birmingham and the Pittsburgh, St. Louis and Fresno Supersites. Transmissometer data should be available for Phoenix, and airport ASOS data may provide useful independent light scattering measurements for some urban areas if it can be obtained in its raw, uncensored form. These data should be used to help identify biases and to provide insight regarding the precision of estimated extinction.

16) In addition to a qualitative discussion of possible sources of uncertainty and variability, are there quantitative methods for addressing uncertainty and variability associated with these assessments that the Panel would recommend?

Although this chapter does present information that pertains to both uncertainty and variability, it could benefit from more structure. For example, this Assessment could be consistent with the Risk and Exposure Assessment (REA) in referencing a framework for dealing with variability and uncertainty, such as the WHO (2008) framework, and explaining as to which tiers of assessment are applied here. Tier 0 is a point estimate based on default values. Tier 1 includes qualitative (but structured) assessment and comparison of sources of uncertainty, including bias and imprecision. Tier 2 includes sensitivity analysis. Tier 3 includes quantitative analysis of uncertainty using probability distributions. For this Assessment, at a minimum a structured Tier 1 approach, supplemented with some Tier 2 applications, would be appropriate.

For models that are calibrated based on empirical data, a statistical estimate of the goodness-of-fit of the model and the standard error of the estimate should be made. The latter can be used as part of a Tier 3 approach to uncertainty analysis under the WHO framework.

17) A number of appendices are provided at the end of this document. Does the Panel agree that this information is useful to retain? Does the Panel agree with the level of detail provided in the body of the report and its organization and distribution throughout the document?

It is useful to retain the appendices, which are needed to support the discussions in the main document. Given the similarities and differences between the currently proposed metrics for a secondary standard, and those considered for a sub-daily fine mass-based urban visibility standard in the last round of PM NAAQS review, we conclude that the level of detail presented in the document is appropriate to adequately inform the Administrator as to the basis for recommending a new secondary standard. Hopefully these appendices can be revised to reflect the changes that would result from considering some of the alternative options CASAC has suggested.

18) Does the Panel agree with the approaches used to simulate just meeting air quality conditions for the current and alternative PM standards? In particular, is use of the

proportional rollback approach appropriate in the context of the urban PM visibility assessment?

It seems unlikely that all PM species would be rolled back by the same proportions in efforts to just attain existing or new primary or secondary PM standards. However, there seems to be no logical basis for suggesting specific alternatives to the proportional rollback model for use in this application. Table 4-7 is an informative display of relevant information which might be expanded for the final visibility risk assessment or in the policy assessment to include other primary PM standard combinations, like 13/30, etc. CASAC has suggested variations in several elements of the secondary PM standard, and it would be informative to see how those revisions – in possible combinations with higher percentiles - might change the relative mixtures of primary PM and secondary light extinction levels across the urban study areas.

The fact that the proposed indicators (and averaging times, levels and forms) for the primary and secondary standards are entirely independent from each other, makes it inherently more difficult to assess particular combinations of primary and secondary NAAQS metrics that are likely under future rollback scenarios. At the same time, the totally independent nature of the primary and secondary indicators and of the health and welfare objectives they address makes it less critical to know the exact future combinations that might result from hypothetical rollback scenarios. The standards can be set independently, using metrics that best reflect the separate health and welfare objectives, regardless of the progress toward one that may (or may not) be made with progress toward the other.

19) To what extent does the Panel consider the presentation of "what if" scenarios for retention of the current secondary PM2.5 NAAQS and consideration of alternative, more protective secondary NAAQS in sections 4.2 and 4.3 to be clearly written with an appropriate level of detail? Do the correlation analyses presented in Appendix D provide sufficient insight into the suitability of alternative indicators based on sub-24 hour averaging periods for PM2.5? Are there additional alternative standard scenarios that should be evaluated?

The presentation of “what if” scenarios for the current standards and alternative new secondary standards is clearly written, with appropriate details conveyed in sections 4.2 and 4.3 and associated appendices. As indicated above, CASAC recommends evaluating several revisions to the proposed secondary standard metrics that will generally tend to dry out the aerosol, somewhat reduce the current East/West differences and result in lower extinction estimates. It would be informative to see the results of some of these “softer metrics” compared to the alternative primary PM standards. Since those changes would essentially screen out many of the worst visibility hours from the regulatory metric, it would also be appropriate to consider combining them with a higher percentile forms – up to and including the 98th percentile.

The correlation analyses presented in Appendix D help show similarities and differences among various sub-daily extinction and fine mass indicators across the different urban areas. The correlations between 4-hour afternoon light extinction and $PM_{2.5}$ are surprisingly poor for several sites (St. Louis and Philadelphia). Presumably the afternoon correlations for these and many other sites will improve if the 95% RH cap is replaced with a 90% RH screen, as recommended. If not, it would be useful to take a closer look at the estimation methods and input data.

1

Enclosure C

Compendium of Preliminary Pre-Meeting Comments

CASAC Particulate Matter Review Panel

Particulate Matter Urban-Focused Visibility Assessment (September 2009)

Ashbaugh Comments (Dr. Lowell Ashbaugh) 2
Cowling Comments (Dr. Ellis Cowling)..... 4
Frey Comments (Dr. H. Christopher Frey)..... 12
Grantz Comments (Dr. David Grantz) 15
Helble Comments (Dr. Joseph Helble)..... 16
Hopke Comments (Dr. P. K. Hopke)..... 18
Malm Comments (Dr. William Malm)..... 20
Poirot Comments (Mr. Rich Poirot)..... 25

Ashbaugh Comments (Dr. Lowell Ashbaugh)

Comments on 1st Draft Urban-Focused Visibility Assessment for PM, September 2009, Lowell Ashbaugh

Charge Questions 12-14:

Chapter 3 – Estimation of Current PM Concentrations and Light Extinction

12) Are the goals articulated in the first paragraph of this chapter achieved in the remainder of the assessment? If not, does the Panel have suggestions for additional assessments that should be done?
13) Are the methods and approaches taken in these assessments, including those for monitor site selection, incomplete data adjustments, and the use of the CMAQ model to augment speciation data, appropriate and is the rationale for their selection clearly articulated?
14) Is the approach used to estimate PRB as described in chapter 3 and Appendix C appropriate?

The goals articulated in the first paragraph of Chapter 3 are largely achieved in the document. There is one addition that may be appropriate - the IMPROVE program has samplers in several urban areas, including some of the 15 cities examined here. The IMPROVE data could be used as a check on the results of the method used here. It's also not exactly correct that there is no systematic monitoring network in place for $PM_{10-2.5}$. The IMPROVE network routinely measures $PM_{10-2.5}$, including at selected urban areas, so the data may be useful for this analysis.
The approach taken to assess urban visibility is sound. Given the relative lack of measurements (except as noted above for IMPROVE) this approach to obtain estimates of hourly daytime visibility is a good one. The method is clearly articulated.
The approach used to estimate PRB is also sound. It might be good to compare the "best" IMPROVE visibility measurements with the PRB results, though. Whenever possible, it serves as a useful check on the modeled results to compare them to appropriate measurements.
I'm uncomfortable with the use of a ratio of $PM_{10-2.5}$ to $PM_{2.5}$ to allocate hourly measurements. Whenever possible, it would be useful to use a daily ratio to perform the allocation instead of a longer time average (monthly or annually). Both PM_{10} and $PM_{2.5}$ change rapidly, but PM_{10} does so especially.
Table 3.2 describes Fresno as being in southern California. Although as a "real" northern Californian, I don't consider Fresno to be in northern California, I can't say it's regarded as being in southern California, either. That distinction normally applies to that part of California south of the Tehachapi Mountains. Fresno is representative of the Central Valley, but not southern California.
The sentence that ends on page 3-11, lines 11-12, seems to be missing something but I can't figure out what it needs to be coherent.
The x-axis labels for Figure 3-4 (both figures) are incorrect. They should run from 01:00:00 am to 12:00:00 pm.
Page 3-18, line 25, should read "principal", not "principle". In the discussion of nitrate sampling on the same page, there should be a note that the fate of ammonium collected on a Teflon filter is not known. In particular, it's known that ammonium nitrate can evaporate from a Teflon filter during collection. It's suspected the ammonium will be lost along with the nitrate, but this has

not been satisfactorily investigated. The same thing is true of ammonium collected on a nylon filter, but in that case it's well documented that the nitrate is retained.

I'm not sure what the discussion of sulfate on page 3-19 is intending to say. What does it mean that the continuous instruments are more like the FRM than the CSN in regards to sulfate? And why is it similar to nitrate, when nitrate is volatile but sulfate is not?

On page 3-24, $PM_{10-2.5}$ was obtained by subtracting $PM_{2.5}$ from PM_{10}. But in setting negative values to zero, you bias the results. There is uncertainty inherent in all PM_{10} and $PM_{2.5}$ measurements. If the measurements are close, there is a probability that the difference will be negative. This possibility should not be ignored.

The sentence on page 3-27, lines 8-10, could be more clearly stated. The phrase "Except for the two Texas and the non-California western urban areas…" is confusing. I believe it would be clearer to say "All urban areas east of the Mississippi, and California and Texas urban areas…" but I defer to the authors to make the sentence easier to understand.

Table 3-6 contains a summary line labeled "*Average*" that is truly an average. But it should be a weighted average, as each urban area has a different number of days with estimates. A weighted average would have 87%, 64%, and 35% instead of 86%, 61%, and 33% for the three CPLs.

On page 3-36, lines 23-25, it could be noted that the extreme haze hours at Fresno are exclusively late fall and winter. Extreme days in Los Angeles span the year. The causes of haze at these two sites are quite distinct; it should not be suggested that they are similar.

Cowling Comments (Dr. Ellis Cowling)

Individual Comments on the Separate Risk-Assessment Document for PM Welfare Effects on Urban-Focused Visibility

Before offering my Individual Comments in response to the 19 Charge Question on the Risk-Assessment Document for PM Welfare Effects on Urban-Focused Visibility, permit me to indicate how pleased I was to read the following statement in Lydia Wegman's transmittal note of September 9:

> **"In addition, on or about September 15, 2009, we plan to release a very preliminary draft of another document, *Policy Assessment for the Review of the Particulate Matter National Ambient Air Quality Standards: Preliminary Draft* (Policy Assessment), to facilitate discussion with the CASAC PM Panel on our ongoing efforts to prepare an external review draft Policy Assessment."**

Chapter 1 – Scope of Visibility Assessment

Charge Question 1:

> **After careful consideration of the evidence provided in the second draft ISA, and in particular the significant body of work that has been conducted by the Regional Planning Organizations under the Regional Haze Rule, (i.e. information on urban and rural PM concentrations and compositions), we have decided to continue to focus this assessment on the PM induced visibility impairment that is occurring in urban areas. What is the Panel's view on this approach? Is the rationale supporting the selected approach clear and appropriate?**

I consider this approach to be very appropriate for the several reasons outlined on page 1-6 of Chapter 1 including: 1) Generally higher ambient air concentrations of PM in urban than in rural areas, 2) Availability of more science-based assessments in urban than in rural areas, and 3) The larger numbers of persons in urban than in rural areas whose enjoyment of decrease visibility impairment would be affected. The approach taken – use of light extinction as the air-quality indicator of concern and the decision to minimize the influence of compounding effects of relative humidity by making compliance measurements in afternoon hours when variation in relative humidity usually are less than in other hours of the day is clearly explained and seems well justified.

Charge Question 2:

> **After further considering the nature of urban versus more remote area PM, and in light of discussions with CASAC at the April 2, 2009 meeting, we have decided not to develop an urban optimized algorithm at this time, but instead to rely on the original IMPROVE algorithm to relate urban PM to local haze (PM light extinction). Is the Panel generally supportive of this approach? Is the rationale supporting this decision appropriate and clearly presented?**

The decision to rely on the original IMPROVE algorithm to relate urban PM concentrations to local haze and light extinction seems well justified and appropriate (as the CASAC PM Panel

indicated earlier). The rationale for this decision is explained clearly in the text of this Urban-Focused Visibility Assessment document.

Charge Question 3:

> **In a change from the planned approach presented in the Scope and Methods Plan, we have decided to conduct a reanalysis of the urban visibility preference studies available at the time of the 2006 PM NAAQS review, rather than conducting new public preference studies since it is highly unlikely that the results of new studies could be completed in time to inform this review. This reanalysis was designed to explore the similarities and differences (comparability) between the current studies and to assess what information could be drawn from these results to inform the selection of visual air quality (VAQ) candidate protection levels (CPLs) to be used in subsequent impact assessments. This reanalysis also includes a recent study by Smith and Howell (2009) for Washington, D.C. which was presented to the CASAC during the public comment phase of the April 2, 2009 meeting and later provided to EPA staff. Does the Panel agree that the information provided by this reanalysis is useful to inform the selection of CPLs? Does the Panel agree that inclusion of the Smith and Howell (2009) is appropriate in both the ISA and Visibility Assessment? To what extent does the Panel consider that the reanalysis of the urban visibility preference studies is clearly and appropriately characterized?**

Yes, I agree with the decisions to reanalyze the urban visibility data and information available at the time of the 2006 PM NAAQS review and to take advantage of the results obtained in the Smith and Howell 2009 study in Washington DC rather than to conduct new public preferences studies. I also agree that the information provided by the reanalysis approach is useful to inform the selection of candidate protection levels (CPLs). But I suggest that the term "candidate protection PM concentrations" (CPPMCs) be used instead of CPLs in order to be clear about the distinction between the air concentrations of PM actually MEASURED at a monitoring site and the air concentrations that might be ALLOWED under a proposed new PM secondary standard.

Charge Question 4:

> **We have chosen to use the range that represents the 50th acceptability criteria across the four cites studied (i.e., the VAQ level that best divides the photographs shown into two groups: those with a VAQ rated as acceptable by the majority of the participants, and those rated not acceptable by the majority of participants) as CPLs to characterize the nature of the impact on urban VAQ associated with current PM levels. Please comment on the clarity and appropriateness of the rationale supporting this decision. Does the Panel have suggestions for alternative ranges to consider?**

This Charge Question mentions "the photographs shown" – which I assumed meant "the photographs shown in this Urban Focused Visibility Assessment document," but when I couldn't find any of the various sets of photographs used by the participants in the various visibility preference studies discussed in this document, I came to realize that the term "photographs shown" must have been intended to refer to "the photographs shown to the participants." Thus without being able to see the photographs in question, I have no personal basis for evaluating the proposal to "use the range that represents the 50th acceptability criteria across the four cities" as opposed to any "other alternative ranges" that might be considered."

Please also note the ambiguity in use of the term "current PM levels" in the statement of Charge Question 4 -- when the purpose of this Charge Question is to ask for CASAC guidance about methods to identify "candidate protection levels." As in my response to Charge Question 3, it is very important to distinguish between air concentrations of PM that are actually MEASURED from those that might be ALLOWED under a proposed new PM secondary standard. Frequent use of the same term -- "level" -- to describe both what is MEASURED and what is ALLOWED is why the alternative terms "measured air concentration" and "allowable air concentration" would avoid confusion in the minds of both CASAC Panelists and some air-quality monitoring personnel.

On the other hand, when I carefully read through the detailed descriptions of how the urban visibility preference studies were conducted in the four cities, I was very impressed with the participant-to-participant consistency and apparent scientific rigor of this approach – including especially the displays of participant preferences shown in Figures 2.2, 2.3, 2.5, 2.9, 2.11, and 2.12 and the thorough discussion on pages 2-24 and 2-25 of alternatives hypotheses to explain the differences in numbers of deciviews found "acceptable" to 50% of the participants in the four cities. Thus, I am convinced that the 50% acceptability criterion is a sound basis for establishing reliable candidate protection levels (CPLs).

Charge Question 5:

A new indicator relating ambient PM to urban VAQ (i.e., PM light extinction) has been evaluated to improve our characterization of the relationship between ambient PM and visibility impairment. Is the Panel generally supportive of this approach? To what extent have we provided an adequate justification for the new indicator used?

Yes, as indicated in my response to Charge Question 2, I believe light extinction is a very appropriate means by which to relate "ambient PM" (concentrations) to urban Visual Air Quality (VAQ). I find the justification provided in this Urban-Focused Visibility Assessment document very adequate as a means by which to relate ambient PM concentration to urban VAQ.

Charge Question 6:

An averaging time of one hour as a practical minimum time period was used in this assessment in recognition that, while the visibility impacts are nearly instantaneous, the urban VAQ does not generally change significantly from minute to minute, but does vary from hour to hour. To what extent does the Panel support this approach? Does the Panel consider the rationale supporting this approach to be clearly and appropriately presented?

The justification provided on page 2-26 for the selection of one hour as a suitable "averaging time" for candidate visibility protection secondary standards seems very adequate to me.

Charge Question 7:

We have chosen to use the 90th and 95th percentile forms in our assessment of alternative secondary (welfare-based) standards. Please comment on the use of these alternative forms.

I agree with the decision discussed on the bottom of page 2-26 that using the 90th and 95th percentile forms in this Urban Focused Visibility Assessment document is appropriate.

Chapter 2 – Urban Visibility Preference Studies

Charge Question 8:

> **To what extent does the Panel support the graphical displays presented in this Chapter? As currently presented, do these figures summarize the assessment results? We have combined data from multiple studies for two locations – British Columbia and Washington D.C. - Abt Associates, 2001; Smith and Howell, 2009 test 1) as presented in Figure 2-14. Does the Panel agree with developing a composite dataset for each of these two urban areas?**

As indicated in my response to Charge Question 4 (above), I am very impressed with the reliability of the graphical display method described in Chapter 2. Yes, the individual figures for each of the four cities appear to me to summarize the assessment results for each city very nicely. Also, the combined display of data for all four cities shown in Figure 2-14 provides an integrated overview of the results in all four cities. Thus, I simply do not understand what is meant by the last part of this Charge Question 8 – "**Does the Panel agree with developing a composite dataset for each of these two urban areas?"**

Charge Question 9:

> **Despite significant differences in study characteristics (e.g., size, location), to what extent does the Panel support combining and comparing the results from the four cities, as shown in Figure 2-14? What is the Panel's view on the clarity and adequacy of the descriptions of the uncertainties and limitations associated with such a combined assessment and the conclusions that can be drawn from the assessments? Please provide comments on additional insights, uncertainties, or caveats that should be considered.**

As indicated in my response to Charge Question 8, I think the combined analysis of results from all four cities shown in Figure 2-14 demonstrates the general efficacy of the methods used Urban Visibility Preference Studies. The variation in number of deciviews that was found "acceptable" in the four cities suggests that a uniform national standard for all urban areas across the whole nation is probably neither practicable or scientifically well-justified at present. It seems to me that some degree of regional and/or individual urban-area flexibility in implementation should be explored. Perhaps based on consensus judgments of the citizens in different urban centers within individual states. Or perhaps recommending only that flexible "targets" rather than fixed "national standards" should be established that would be designed to facilitate and encourage improvement in urban visibility progressively over time – for example, by state and/or federal requirements and/or incentives that would encourage individual urban areas to design visibility-improvement programs that will decrease the number of days per year that visibility is judged "unacceptable."

Charge Question 10:

> **We have used the combined results presented in this chapter to develop a range of CPLs that are used in subsequent steps of the assessment. To what extent does the Panel support the range of CPLs used and the justification provided for selecting this range? Does the Panel recommend consideration of any alternative approaches or criteria for selecting CPLs?**

The complexity of the analysis described in Chapters 2 and 3 and the variability by region across our country, and the relative lack of consistent data relating directly to light extinction and the

limitations of having to depend on modeling estimates all combine to challenge my own meager understanding of some if not all of these complicating factors. Maybe as much because of- and also in spite of- the bewilderment that I experienced in reading Chapter 3, I believe that EPA staff have developed a range of CPLs that appear very reasonable to me. At the same time, I look forward eagerly to hearing the considered professional judgments of two of my fellow CASAC PM Panel members whose experience and capacity for making informed policy-relevant judgments under conditions of substantial measurement uncertainty is substantially greater than my own – Bill Malm and Rich Poirot.

Charge Question 11:

> **Overall, we consider this assessment useful for providing information for the design of future urban visibility preference studies. Does the Panel support this conclusion and does the Panel have any recommendations for changes that could be made in the discussions of this information to enhance its usefulness for this purpose?**

I agree that the assessment provided in this document is very useful for the design of future urban visibility preference studies and hope that EPA and the various air-quality-management units of several states will join with interested stakeholders to make persistent efforts to ensure that many more such studies are undertaken in the future.

Chapter 3 – Estimation of Current PM Concentrations and Light Extinction

Charge Question 12:

> **Are the goals articulated in the first paragraph of this chapter achieved in the remainder of the assessment? If not, does the Panel have suggestions for additional assessments that should be done?**

Yes, I believe that the goals set out in the first paragraph of Chapter 3 were achieved very substantially. But once again, for many of the same reasons stated above in my response to Charge Question 10, I look forward to hearing the much better informed professional judgments of my colleagues on this PM NAAQS review Panel – especially Bill Malm and Rich Poirot.

Charge Question 13:

> **Are the methods and approaches taken in these assessments, including those for monitor site selection, incomplete data adjustments, and the use of the CMAQ model to augment speciation data, appropriate and is the rationale for their selection clearly articulated?**

Yes, I agree that the methods and approaches taken in these assessments including those for monitor site selection, incomplete data adjustments, and the use of the CMAQ model to augment speciation data are appropriate and that the rationale for their selection was described very adequately.

Charge Question 14:

> **Is the approach used to estimate PRB as described in chapter 3 and Appendix C appropriate?**

Yes, I believe that the approach used in estimating policy-relevant background as described on page 3-37 and in Appendix C was appropriate.

Charge Question 15:

We consider the results generated by these analyses to be reasonable based on PM composition and relative humidity data. Does the Panel agree? Are there other tests of reasonableness that could be applied?

Once again, as indicated in my response to Charge Questions 10 and 12 (above), I believe that EPA staff have done a very good job of dealing with the issues of variation in PM composition and the complicating influence of variation in relative humidity.

Charge Question 16:

In addition to a qualitative discussion of possible sources of uncertainty and variability, are there quantitative methods for addressing uncertainty and variability associated with these assessments that the Panel would recommend?

My experience in dealing with the issues of uncertainty and variability in both PM monitoring and visibility measurements is not sufficient for me to provide useful suggestions about quantitative methods for dealing with the issues of uncertainty and variability in these assessments.

Charge Question 17:

A number of appendices are provided at the end of this document. Does the Panel agree that this information is useful to retain? Does the Panel agree with the level of detail provided in the body of the report and its organization and distribution throughout the document?

My professional judgment is that the information in the appendixes is useful and should be retained. I also agree that the amount of detail in the body of this report and its organization and distribution throughout the document constitute a very good job well done – especially considering the complexity and limitations of available scientific data and information.

Chapter 4 – Total Light Extinction Under "What If" Conditions of Just Meeting Specific Alternative Secondary NAAQS

Charge Question 18:

> **Does the Panel agree with the approaches used to simulate just meeting air quality conditions for the current and alternative PM standards? In particular, is use of the proportional rollback approach appropriate in the context of the urban PM visibility assessment?**

Yes, I agree with the approaches used to simulate just meeting air quality conditions for the alternative light-extinction based PM standards that are discussed in Chapter 4. I also believe that it is almost always sensible and useful to consider a proportional rollback approach in the context of any air-pollution management systems – including urban visibility assessments.

But I am puzzled about how to consider these approaches in the context of the current PM secondary standard which is based on mass of PM rather than on light extinction. Perhaps my confusion about this matter will be decreased during our up-coming CASAC PM NAAQS meeting on October 5 and 6, and then I can reconsider this Charge Question and my response to it when I better understand the transition that will occur if and when we change from the current identical mass-based primary and secondary standards for PM to one or more alternative secondary standards that will be based on light-extinction rather than on mass of PM..

Charge Question 19:

> **To what extent does the Panel consider the presentation of "what if" scenarios for retention of the current secondary $PM_{2.5}$ NAAQS and consideration of alternative, more protective secondary NAAQS in sections 4.2 and 4.3 to be clearly written with an appropriate level of detail? Do the correlation analyses presented in Appendix D provide sufficient insight into the suitability of alternative indicators based on sub-24 hour averaging periods for $PM_{2.5}$? Are there additional alternative standard scenarios that should be evaluated?**

As indicated in my response to Charge Question 18, I am puzzled about how to consider the presentation of "What if" scenarios in the context of the current PM secondary standard which is based on mass of PM rather than on light extinction. Perhaps my confusion about this matter will be decreased during our up-coming CASAC PM NAAQS meeting on October 5 and 6, and then I can reconsider this Charge Question and my response to it when I better understand the transition that will occur if and when we change from the current identical mass-based primary and secondary standards for PM to one or more alternative secondary standards that will be based on light-extinction rather than on mass of PM.

Permit me also to call attention to my response to Charge Question 9 (above) in which I have raised the issue of having flexible "targets" rather than fixed "national standards" for improvement of urban visibility:

> "It seems to me that some degree of regional and/or individual urban-area flexibility in implementation should be explored. Perhaps based on consensus judgments of the citizens in different urban centers within individual states. Or perhaps recommending only that flexible "targets" rather than fixed "national standards" should be established that would be designed to facilitate improvement in urban visibility progressively over time – for example, by state and/or federal requirements and/or incentives that would

encourage individual urban areas to design visibility-improvement programs that will decrease the number of days per year that visibility is judged "unacceptable."

Frey Comments (Dr. H. Christopher Frey)

I was asked to respond to Charge Questions 15, 16, and 17.

Charge Question 15: *We consider the results generated by these analyses to be reasonable based on PM composition and relative humidity data. Does the Panel agree? Are there other tests of reasonableness that could be applied?*

Response: EPA has undertaken a reasonable review of available data and development of databases from which light extinction has been estimated. The choice of the estimation equation is explained. The data and equation are applied to case study examples for numerous urban areas. Thus, the assessment takes into account variability in: (a) the daily mass concentration of $PM_{2.5}$ and coarse PM ($PM_{10-2.5}$), (b) speciation; and (c) relative humidity. The document discusses the geographic variability in speciation and provides plausible explanations for inter-regional variability. The capabilities and limitations of the existing monitoring networks are appropriately identified. The hierarchical approach for use of FRM data, FEM data, and estimates from chemical transport models is reasonable. The development of diurnal patterns of $PM_{2.5}$ components is reasonable. The summaries of data are useful and are interpreted reasonably. For example, the distribution of hourly estimated light extinction and the discussion of the patterns in these estimates is very useful. The discussion of the implications of observations of relative humidity of greater than 95% is useful and appropriately developed. As pointed out, such high humidity may often be associated with precipitation or fog events in which there is not an expectation of high visibility.

In terms of the data regarding PM sizes and their components, EPA has considered the appropriate measurement and modeling data sets. With respect to the discussion of humidity levels, perhaps there could be comparison to precipitation measurement data to quantify or at least provide qualitative insight into the proportion of relative humidity data above 95 percent that are associated with precipitation events. This would further strengthen what may already be an adequate argument for capping RH at 95 percent when inputting data to the IMPROVE equation.

Charge Question 16: *In addition to a qualitative discussion of possible sources of uncertainty and variability, are there quantitative methods for addressing uncertainty and variability associated with these assessments that the Panel would recommend?*

Response: Although this chapter does present information that pertains to both uncertainty and variability, it could benefit from more structure. For example, this Assessment could be consistent with the Risk and Exposure Assessment (REA) in referencing a framework for dealing with variability and uncertainty, such as the WHO (2008) framework, and explaining as to which tiers of assessment are applied here. Tier 0 is a point estimate based on default values. Tier 1 includes qualitative (but structured) assessment and comparison of sources of uncertainty. Tier 2 includes sensitivity analysis. Tier 3 includes quantitative analysis of uncertainty using probability distributions. For this Assessment, at a minimum a structured Tier 1 approach, supplemented with some Tier 2 applications, would be appropriate.

A good example of how to implement the Tier 1 approach is illustrated in Table 3-13 (p. 79) of *Risk Assessment to Support the Review of the PM Primary National Ambient Air Quality Standards, External Review Draft* (September 2009). A similar tab le in this Assessment would be a useful way to summarize and help synthesize the implications of the various sources of uncertainty. Such a table would contain an identifier and description of the source of uncertainty, a statement of the direction (high, low, both, or none) of the response of the assessment endpoint to uncertainty in the input, an assessment of the knowledge base, and a commentary that provides rationale for the assessment. In addition, there should be some discussion of relationships (if any) among pairwise combinations of sources of uncertainty. The qualitative discussion should conclude with an assessment of which sources of uncertainty are the most significant, and whether or how such sources of uncertainty would bias the answer or lead to imprecision in the answer.

In some cases, a Tier 2 sensitivity analysis would be useful. It would be helpful to present a simple evaluation of how the estimated light extinction from the IMPROVE equation varies based on ranges of values for each input that are representative of observed values of PM component and mass concentration and relative humidity. For example, the results in Figures 3-12 to 3-19 generally (with some exceptions, of course) imply that organic carbon, sulfate, and nitrate are often the three most important components that affect light extinction. This result is not entirely obvious just by inspection of the Equation on page 3-16 because the sensitivity of the estimated light extinction depends on two factors: (a) on the functional form and the coefficient for each term; and (b) the typical values (and variability) for the inputs.

Furthermore, there is considerable discussion in the text regarding some of the uncertainties associated with apportioning $PM_{2.5}$ to the mass of components. However, there seems to be no analysis of whether or by how much these potential errors would make a difference when estimating light extinction. Can statements be made as to how much these factors might introduce errors in terms of either relative (%) or absolute (ug/m^3) bias, imprecision, or both? Such statements could be translated into a sensitivity analysis for a few representative cases to demonstrate by how much would the estimate of light extinction vary.

The difficulty in estimating the coarse PM concentration, particularly when PM_{10} and $PM_{2.5}$ monitors were not sited concurrently, might lead to more uncertainty in light extinction estimates for some regions than others. For example, Phoenix appears to have a larger contribution of coarse PM to light extinction than many other areas, although on a given day coarse PM appears to be the dominant contributor in other locations such as Los Angeles and St. Louis. Since coarse PM is expected to have more geographic variability than fine PM because of higher deposition rates and because of the mode of its formation, it would be important to at least emphasize as to whether the coarse PM data for these three cities are based on collocated monitors. If not, then some additional qualitative discussion, and perhaps sensitivity analysis, could provide insight regarding the robustness of the estimates of light extinction.

Charge Question 17: *A number of appendices are provided at the end of this document. Does the Panel agree that this information is useful to retain? Does the Panel agree with the level of detail provided in the body of the report and its organization and distribution throughout the document?*

Response: The level of detail in the body of the report is appropriate. It is appropriate that additional details are in the appendices.

Grantz Comments (Dr. David Grantz)

Grantz Individual Comments:

Charge Question 1.

The rationale for pursuing a Secondary PM Standard based on daylight visual impairment in urban areas is thoroughly and clearly laid out in this document. The approaches, the reliance on computed visibility, and the concept of preference studies are appropriate. The authors have prepared a very complete, readable and informative document. Given the momentum in this direction, it is clear that this round of PM assessment is likely to proceed along these lines. However, I am not convinced that this is the best path.

I do not find the arguments compelling that a daytime only standard is appropriate. Is night-time dumping of PM acceptable with respect to welfare?

I am not convinced that an urban standard is appropriate. Is it ok to dump PM to the atmosphere if it will diffuse to rural areas relatively quickly?

I am particularly concerned that visibility should be taken to represent welfare effects. It does have the distinct advantage of being amenable to calculation from existing mass measurements of PM, in contrast to the potentially more important impacts on materials, ecosystems, and climate. However, the current near-total focus on urban visibility may come to constrict our ability to identify or assess other endpoints, including by redirecting research funds to the now policy relevant visibility issue.

Charge Question 9.

I am not particularly comfortable with the combined analysis of the preference studies from the four cities. The authors have presented the uncertainties and analytical issues fairly well. However, despite the consideration of four different hypotheses to explain the different VAQ preferences in east and west, I think the actual reason was missed. It is a subset of Hypothesis 4, but more closely related to individual experience after some time in a location. What differs is personal expectations, not differences in background, lighting etc. It is not reasonable to expect that people in eastern cities who rarely see 20 miles would express a preference for 40 miles, whereas in the west this visual range is not unusual and might be expected. As a result, the data are inherently not-combinable. They define different things. Some manipulation of ambient RH could bring this all back to mass, and make the data compatible, but this would not really be a visibility standard.

Helble Comments (Dr. Joseph Helble)

Question 2: The decision to use the IMPROVE algorithm to relate urban PM to local haze is reasonable. Given the uncertainties in the underlying data, particularly PM composition, needed to model visibility, the rationale for utilizing the existing algorithm is appropriate and clearly articulated by the Agency.

Questions 5 – 6:

5) A new indicator relating ambient PM to urban VAQ (i.e., PM light extinction) has been evaluated to improve our characterization of the relationship between ambient PM and visibility impairment. Is the Panel generally supportive of this approach? To what extent have we provided an adequate justification for the new indicator used?

The use of visual air quality as an indicator of urban VAQ is reasonable, and adequate justification is provided. Some discussion of the shortcomings of this approach (resulting from this measurement being a local estimate of a quantity that is integrated over km length scales) would have been helpful, but it is acknowledged that point PM mass measurements are similarly constrained.

6) An averaging time of one hour as a practical minimum time period was used in this assessment in recognition that, while the visibility impacts are nearly instantaneous, the urban VAQ does not generally change significantly from minute to minute, but does vary from hour to hour. To what extent does the Panel support this approach? Does the Panel consider the rationale supporting this approach to be clearly and appropriately presented?

One hour averaging is a reasonable compromise.

Chapter 4 – Total Light Extinction Under "What If" Conditions

Questions 18 and 19: Proportional rollback, what-if scenarios

The approaches EPA has used are reasonable. Given the uncertainty in calculations of light extinction by PM, proportional rollback is reasonable.

Regarding Sections 4.2 and 4.3 – the information is clearly presented, although Figure 4.1 is difficult to read in print. The on-line version, with the figure expanded at least 2x, is necessary for seeing detail.

Figure 4-1: Comparison of the different scenarios would benefit from a single figure comparing these for a single city.

Chapter 3 – Estimate of Current PM Concentrations and Light Extinction

Questions 12 and 13: Goals, Methods, Approaches:

The goal of Chapter 3 was to examine urban study area light extinction and develop an improved understanding of the underlying causes and patterns. The stated goal of Chapter 4 was to

examine "what if" scenarios. The goal of Appendix D, to examine potential alternative indicators for light extinction.

Chapters 3 and 4 meet the goals articulated in the first two sentences above. Chapter 3 in particular provides a thorough assessment of the limitations of existing data and the uncertainties related to an assessment of visibility (light extinction).

Appendix D presents data on the relationship between various PM measures and light extinction. The data are presented clearly, but there is no discussion of the findings. Interpretation would have been helpful.

It appears that EPA generalized each "season" so that each location was defined to have the same number of daylight hours. It is not clear what benefit this provides, and the suggestion (footnote 23) that this may be eliminated in the final version of this assessment seems appropriate.

The sentence beginning on the bottom of page 3-25 (line 28): "Persons…" seems speculative. Are there studies to support this? I suspect that those in foggy locations would *value* clear visibility, even if they do not expect it.

Line 31, page 3-36: not clear what is meant by "..carbonaceous-caused extreme hours…" Presumably this refers to periods of extremely high light extinction caused by a higher than normal concentration of carbon-containing PM.

Typo - Appendix D page D-1 refers to CMAQ profiles in section 3.2.4. This report does not contain a section 3.2.4.

Question 14 – PRB

The approach used to estimate PRB is reasonable.

Other comments

1. delete "that will be developed" from line 8, page 1-2
2. line 13, p. 1-7, rational should be rationale
3. line 32 p 1-10, should read urban areas (plural)

Hopke Comments (Dr. P. K. Hopke)

Comments by P.K. Hopke on the Particulate Matter Urban-Focussed Visibility Assessment

This assessment generally makes appropriate use of the available ambient aerosol data. There are some issues that need to be addressed to improve the input values, but the approach taken in this assessment are reasonable and appropriate to the task at hand.

On page 3-4, line 9, It says "most accurate method" There is no basis for detemining "accuracy" with respect to any ambient mass monitor system. Precision can be assessed, but not accuracy.

There seems to be a misconception that the FRM measures ambient mass concentrations related to light scattering. This is clearly incorrect. We know that the FRM loses significant quantities of both nitrate and semivolatile organic matter so that it will generally underestimate the mass concentration that would be present to degrade visibility. Thus, the methodology that adjusts the continuous measurement of PM mass to match the FRM are misguided. The unmodified FDMS measurements are probably the best estimates we have of the actual airborne PM mass (although we really have no good way to verify this hypothesis). It will generally produce values greater than the FRM so that scaling the continuous measurements down to the FRM value will result in overestimating the effect of mass on light extinction.

They have probably done as well as they can with the $PM_{2.5-10}$ values since EPA has been unwilling to develop an effective monitoring system for this size class. The results presented here continue to demonstrate the problem with the difference method and it is unfortunate that they have not built impactor based systems to actually measure this size fraction mass.

The choice of data for mass and composition are reasonable. In terms of uncertainty and variability, they have again done a reasonable job.

They should retain the appendices.

5) A new indicator relating ambient PM to urban VAQ (i.e., PM light extinction) has been evaluated to improve our characterization of the relationship between ambient PM and visibility impairment. Is the Panel generally supportive of this approach? To what extent have we provided an adequate justification for the new indicator used?

Yes, the public perception of poor air quality when visibility is diminished. Light extinction is an easily measured quantity that can be related in a reasonable way to observed visibility so that it provides an objective measure of VAQ. Thus, the framework for adopting an extinction measurement as the basis for a new indicator have been appropriately presented.

6) An averaging time of one hour as a practical minimum time period was used in this assessment in recognition that, while the visibility impacts are nearly instantaneous, the urban VAQ does not generally change significantly from minute to minute, but does vary from hour to hour. To what extent does the Panel support this approach? Does the Panel consider the rationale supporting this

approach to be clearly and appropriately presented?

At this time, the idea of an hourly averaging period is reasonable and the presentation in the document is credible. In the absence of very much data, it is hard to fully assess the temporal variability, but it is likely that in most circumstances, such an averaging period would be reasonable and obviously then matches the time frame for which the commonly measured pollutant variables are measured.

15) We consider the results generated by these analyses to be reasonable based on PM composition and relative humidity data. Does the Panel agree? Are there other tests of reasonableness that could be applied?

The results presented in this assessment are reasonable given the data that are available and the current state of understanding of the effect of composition on extinction. In the absence of measured hourly compositional data, it is hard to envision other tests of reasonableness.

16) In addition to a qualitative discussion of possible sources of uncertainty and variability, are there quantitative methods for addressing uncertainty and variability associated with these assessments that the Panel would recommend?

Although this chapter does present information that pertains to both uncertainty and variability, it could benefit from more structure. For example, this Assessment could be consistent with the Risk and Exposure Assessment (REA) in referencing a framework for dealing with variability and uncertainty, such as the WHO (2008) framework, and explaining as to which tiers of assessment are applied here. Tier 0 is a point estimate based on default values. Tier 1 includes qualitative (but structured) assessment and comparison of sources of uncertainty. Tier 2 includes sensitivity analysis. Tier 3 includes quantitative analysis of uncertainty using probability distributions. For this Assessment, at a minimum a structured Tier 1 approach, supplemented with some Tier 2 applications, would be appropriate.

17) A number of appendices are provided at the end of this document. Does the Panel agree that this information is useful to retain? Does the Panel agree with the level of detail provided in the body of the report and its organization and distribution throughout the document?

It is useful to retain the appendices. Without them, it would be difficult to follow the arguments presented in the document in detail. Given that there were questions regarding the setting of an urban visibility standard in the last round of standard review, we feel that the level of detail presented in the document is appropriate to adequate inform the Administrator as to the basis for recommending a new secondary standard.

Malm Comments (Dr. William Malm)

1) After careful consideration of the evidence provided in the second draft ISA, and in particular the significant body of work that has been conducted by the Regional Planning Organizations under the Regional Haze Rule, (i.e., information on urban and rural PM concentrations and compositions), we have decided to continue to focus this assessment on the PM induced visibility impairment that is occurring in urban areas. What is the Panel's view on this approach? Is the rationale supporting the selected approach clear and appropriate?

I believe the approach is appropriate but the "assessment" document could be more clearly presented. The rationale given for the selected approach is appropriate. Some issues concerning clarity are outlined below.
On page 3-18 it's stated that the mass scattering multipliers are different for each of the species in the IMPROVE algorithm – they are the same for sulfates and nitrates. The discussion under nitrates on the same page is somewhat misleading. The FRM sampler provides an estimate of gravimetric mass that can be severely underestimated because of nitrate volatilization; where nitrates are a significant fraction of $PM_{2.5}$, the underestimation may be greater than 50%. Some range of uncertainty of error associated with the FRM sampler should be discussed. How can nitrate not be fully neutralized? (Not like sulfate.) Also, nitrate may be in the form of calcium or sodium nitrate, since in some cases measured $PM_{2.5}$ nitrate is the fine tail of the coarse mode.
Under "sulfate" on page 3-19 the statement is made "…continuous $PM_{2.5}$ instruments can be assumed to be more like FRM…" What is the justification for this statement? I don't believe it to be true!
Page 3-19: Some reference is made to the SANDWICH approach; this approach should be explained at some minimal level. Figure 3.5 outlines the sequence of steps to estimate hourly $PM_{2.5}$ components and total light extinction from modeling results. Some effort should be made to assess the potential uncertainty in the estimated extinction after all the normalizations and adjustments. The approach may very well be reasonable and the best that can be done under the circumstances, but the renormalizations will tend to overestimate the contribution of some species to extinction while underestimating others. A careful examination of potential biases should be explored. There are some temporally continuous speciated datasets "out there" that can be used to compare to the approach outlined in Figure 3.5: the SEARCH dataset (at least Birmingham and Atlanta), all the supersites, certainly Pittsburgh and Fresno.
Page 3.24: It is stated that the contribution of coarse mass to extinction is low compared to $PM_{2.5}$. How low? Might want to include a table showing $PM_{2.5}$ extinction compared to PM_{10} - $PM_{2.5}$.
Page 3.25: When did you start counting the number of minutes the sun was up for a full hour of sunlight, when the sun was just visible or when the whole disk was above the horizon? It wasn't entirely clear as to how the 90th or 95th percentile extinctions values were used to pick design values. I assume that the highest 1-hour extinction value for each day was selected and then the 90th or 95th percentile was selected for that distribution. It seems that RH was capped at 95% and then these values went into the average. If the RH was 99%, it was set to 95% and then the scattering was calculated based on the commensurate modeled concentrations but with 95% RH. What was done needs to be clearly stated. As written on page 4.3, it seems that this one value for each of 3 years was averaged to get a resulting design value. However, I assume that all values above the 90th or 95th percentile were averaged, and these averages over years were

further averaged to get the final design value. Need to clarify! How were the PRB values applied? Yearly averages and then subtracted, or averaged over 3 years and then subtracted? Again, it should be made explicitly clear. Should discuss the uncertainty or biases associated with a linear rollback. All species will not decrease at the same rate. Might do some sensitivity analysis to see what it means to have rollback that affects mostly sulfate versus nitrates versus organics. The resulting extinction could be quite different, depending on the scheme one uses.

2) After further considering the nature of urban versus more remote area PM, and in light of discussions with CASAC at the April 2,2009, meeting, we have decided not to develop an urban optimized algorithm at this time, but instead to rely on the original IMPROVE algorithm to relate urban PM to local haze (PM light extinction). ls the Panel generally supportive of this approach? Is the rationale supporting this decision appropriate and clearly presented?

I support this approach. In lieu of the uncertainty of R_{oc} factors for urban environments, other biases in estimating species concentrations in urban areas, and lack of optical measurements in urban areas, the use of an IMPROVE-type extinction algorithm that adjusts mass scattering efficiencies on the basis of mass concentrations would be inappropriate. However, the simpler original IMPROVE equation should be modified to include seasalt, and some thought should be given to what the appropriate R_{oc} factor might be. Certainly 1.8, as used in the IMPROVE equation, may be too high. Something around 1.2–1.4 may be more appropriate.

3) In a change from the planned approach presented in the Scope and Methods Plan, we have decided to conduct a reanalysis of the urban visibility preference studies available at the time of the 2006 PM NMQS review, rather than conducting new public preference studies since it is highly unlikely that the results of new studies could be completed in time to inform this review. This reanalysis was designed to explore the similarities and differences (comparability) between the current studies and to assess what information could be drawn from these results to inform the selection of visual air quality (VAQ) candidate protection levels (CPLs) to be used in subsequent impact assessments. This reanalysis also includes a recent study by Smith and Howell (2009) for Washington, D.C., which was presented to the CASAC during the public comment phase of the April 2, 2009 meeting and later provided to EPA staff. Does the Panel agree that the information provided by this reanalysis is useful to inform the selection of CPLs? Does the Panel agree that inclusion of the Smith and Howell (2009) is appropriate in both the ISA and Visibility Assessment? To what extent does the Panel consider that the reanalysis of the urban visibility preference studies is clearly and appropriately characterized?
I support the reanalysis effort, the selection of CPLs, and the inclusion of the Smith and Howell data. The analysis could be done in a more elegant and statistically defensible way. A logistic distribution could be fitted to the data and percentile values extracted from the curve fit, along with associated confidence intervals. Then one would have some idea as to whether the values derived from various studies are indeed statistically different from each other.

4) We have chosen to use the range that represents the 50th acceptability criteria across the four cites studied (i.e., the VAQ level that best divides the photographs shown into two groups: those with a VAQ rated as acceptable by the majority of the participants, and those rated not acceptable by the majority of participants) as CPLs to characterize the nature of the impact on urban VAQ associated with current PM levels. Please comment on the clarity and

appropriateness of the rationale supporting this decision. Does the Panel have suggestions for alternative ranges to consider?
The presentation is clear. I do not have strong feelings about what percentile should be selected. I guess I would choose, as has been done in various perception threshold studies, something like the 90th percentile. Here again, it would be nice to have a logistic model developed that would allow a selection of any percentile with a defined uncertainty. Selection of percentile levels by "eye" as was done in this assessment is certainly less elegant but not necessarily inappropriate.

5) A new indicator relating ambient PM to urban VAQ (i.e., PM light extinction) has been evaluated to improve our characterization of the relationship between ambient PM and visibility impairment. Is the Panel generally supportive of this approach? To what extent have we provided an adequate justification for the new indicator used?
Justification for using extinction has been provided, and because it is more representative of "visibility", it certainly is more appropriate than a mass design level.

6) An averaging time of one hour as a practical minimum time period was used in this assessment in recognition that, while the visibility impacts are nearly instantaneous, the urban VAQ does not generally change significantly from minute to minute, but does vary from hour to hour. To what extent does the Panel support this approach? Does the Panel consider the rationale supporting this approach to be clearly and appropriately presented?
*It has been demonstrated that significant changes in mass concentrations can take place on temporal increments of minutes. However, averaging these changes over distances of many kilometers may average out to small changes in "visibility" or path average extinction. Some quantitative justification should be given as to why an hour is the smallest time increment that should be considered. What are the standard deviations of hourly averages of various aerosol species? These data again are available at some urban areas and certainly for the "super" sites.*7) We have chosen to use the 90th and 95th percentile forms in our assessment of alternative secondary (welfare-based) standards. Please comment on the use of these alternative forms.

It is my understanding that it is the 90th or 95th percentile of the max hour on each day is what has been considered. I would suggest that this approach be compared to just selecting the 90th or 95th percentile extinction for all days (only daylight hours) without concern for daily maximums. What are the advantages or disadvantages of one approach over the other? What are the potential effects of these choices on emission control strategies? Will they be any different? At first glance at Figures 3.12–3.19, it seems that the relative mix of aerosol contribution to extinction is the same on all 90th percentile max extinction days, so maybe it doesn't make much difference which strategy is taken for calculating the design extinction values. However, I would suggest exploring a number of alternative strategies and seeing how the positives and negatives compare and contrast for each approach.

8) To what extent does the Panel support the graphical displays presented in this chapter? As currently presented, do these figures clearly summarize the assessment results? We have combined data from multiple studies for two locations -British Columbia and Washington, D.C. - (Abt Associates, 2001; Smith and Howell, 2009 test 1) as presented in Figure 2-14. Does the Panel agree with developing a composite dataset for each of these two urban areas?

I think you can combine the dataset for a "composite" analysis. However, I think you should develop a logistic model as discussed above.

9) Despite significant differences in study characteristics (e.g., size, location), to what extent does the Panel support combining and comparing the results from the four cities, as shown in Figure 2.14? What is the Panel's view on the clarity and adequacy of the descriptions of the uncertainties and limitations associated with such a combined assessment and the conclusions that can be drawn from the assessments? Please provide comments on additional insights, uncertainties, or caveats that should be considered.
See comments above and comments under question 3.

10) We have used the combined results presented in this chapter to develop a range of CPLs that are used in subsequent steps of the assessment. To what extent does the Panel support the range of CPLs used and the justification provided for selecting this range? Does the Panel recommend consideration of any alternative approaches or criteria for selecting CPLs?

See comments above. Another issue not discussed above is the way the extinction is calculated as a function of RH. I would suggest an upper limit of 90% RH. The uncertainties of RH measurements above 90% are large, and furthermore, with the RH as high as 95%, one could very well have "wisps" of clouds in a sight path, which isn't accounted for in the extinction algorithm. Also, when the RH is above the threshold value for an hour, that hour (or increment of time used to do the extinction calculation) should be excluded from any statistic used to calculate a design value. I believe the current estimates are based on setting the RH to a capped value and including an estimate for extinction at the capped RH value in the average or statistic. If the RH is 100% and it is raining or foggy, aerosols are typically not a significant contributor to visibility reduction.

11) Overall, we consider this assessment useful for providing information for the design of future urban visibility preference studies. Does the Panel support this conclusion and does the Panel have any recommendations for changes that could be made in the discussions of this information to enhance its usefulness for this purpose?
I think that some discussion should be given to the potential role of aerosols on sky color and visibility of clouds, which are both integral to good visibility in an urban setting.

12) Are the goals articulated in the first paragraph of this chapter achieved in the remainder of the assessment? If not, does the Panel have suggestions for additional assessments that should be done?
Goals are well articulated. See comments under question 1. Certainly, model results of speciated aerosol concentrations should be compared to available monitoring data, such as the SEARCH and supersite datasets.

13) Are the methods and approaches taken in these assessments, including those for monitor site selection, incomplete data adjustments, and the use of the CMAQ model to augment speciation data, appropriate and is the rationale for their selection clearly articulated?
See comments above. If you are going to include a discussion of monitoring site selection, then it seems that some discussion of requirements for monitor design should be included. For

instance, how will an ambient scattering measurement be made (as opposed to heating the aerosol before the scattering measurement is made), will the nephelometer have a size-selective inlet, how will coarse particle scattering be measured or estimated, how will ambient absorption be measured, and so forth?. I would suggest that these issues are as or more important than site selection. Either site selection issues should be left out of this document, or general monitoring equipment design criteria should be included.

14) Is the approach used to estimate PRB as described in chapter 3 and Appendix C appropriate?
Yes – see comments above as to how averages were calculated.

15) We consider the results generated by these analyses to be reasonable based on PM composition and relative humidity data. Does the Panel agree? Are there other tests of reasonableness that could be applied?
May or may not be reasonable; compare modeled results to available datasets.

16) In addition to a qualitative discussion of possible sources of uncertainty and variability, are there quantitative methods for addressing uncertainty and variability associated with these assessments that the Panel would recommend?
Compare to available datasets, and then a real handle on uncertainty can be achieved.

17) A number of appendices are provided at the end of this document. Does the Panel agree that this information is useful to retain? Does the Panel agree with the level of detail provided in the body of the report and its organization and distribution throughout the document?
The information should be retained.

18) Does the Panel agree with the approaches used to simulate just meeting air quality conditions for the current and alternative PM standards? In particular, is use of the proportional rollback approach appropriate in the context of the urban PM visibility assessment?
See comments above.

19) To what extent does the Panel consider the presentation of "what if" scenarios for retention of the current secondary $PM_{2.5}$ NAAQS and consideration of alternative, more protective secondary NAAQS in sections 4.2 and 4.3 to be clearly written with an appropriate level of detail? Do the correlation analyses presented in Appendix D provide sufficient insight into the suitability of alternative indicators based on sub-24 hour averaging periods for $PM_{2.5}$? Are there additional alternative standard scenarios that should be evaluated?
See discussion above under question 10. Some alternative approaches should be considered. Picking a 90th or 95th percentile, using all days, as opposed to 1-hour maximum extinction on each day, should be considered. Set the upper RH value to 90%, and any hour with RH above that value should not be included in the design value.

Poirot Comments (Mr. Rich Poirot)

General Comments on the September, 2009 PM Visibility Risk Assessment:
Compliments to the authors – this is an excellent assessment in all respects. Its thoughtfully conceived, well organized and very clearly written. There is also a substantial amount of underlying number crunching and analysis for which I can appreciate many of the difficulties, and am most impressed by how much work you did in such a short time. The complex analyses also required a number of important "decision points" where options were considered and specific choices were made. In all cases, these appear to be logically reasoned and clearly described, such that the reader can follow every step of the analysis, with additional details provided in the appendix. I agree with most, but not all of these choices (though I don't disagree strongly with any). This is an outstanding first draft and should make a very sound basis for recommending a range of secondary standards to the Administrator.

While I concur that a sub-daily PM light extinction indicator could be a very appropriate and effective regulatory metric, I also thought a sub-daily $PM_{2.5}$ mass indicator - proposed in the last PM NAAQS review cycle - could have been effective as well, and that the difference between the two approaches is primarily due to aerosol water content. I think each approach has advantages and disadvantages over the other, and that there is a lot of potentially productive "middle ground" between the "all dry" and "all wet" approaches. For example, the Denver standard uses a PM light extinction indicator but as a daily maximum 4-hour average (rather than 1-hour maximum), constrained to the 8-hour period between 8 AM and 4 PM (not all daylight hours) and limited to hours when RH is less than 70% (not <95%). This is essentially a dryer (and for a given level and percentile form, less stringent) version of what staff is currently proposing, but it could still be an effective basis for a secondary PM NAAQS (accompanied by a higher percentile).

Some specific alternatives that I would like to see considered include:

- Use a RH limit as a screen (don't include hours that exceed it) rather than a cap (converting values above the cap to the cap value). One of the important effects of the screen is to eliminate times with precipitation or fog. A cap doesn't do this.
- Lower the RH screen from 95% to 90%. For similar reasons as above, and also because of the extreme steepness of the f(RH) curve in this area, where a small error in the RH measurement, or in its temporal or spatial representativeness, could make a huge difference in the results. This lower screen will miss some of the worst visibility conditions, but there will be better confidence that the visibility levels in the remaining hours are affected by pollution and not the weather.
- Consider alternatives to the single worst daylight hour in a day as the regulatory metric. Possibilities include use of the second worst hour of the day, use of 2 to 4-hour averages, narrow the "daylight window" to exclude the first and last daylight hours (or keep a constant 8-hour window as they do in Denver), calculate the percentile based on all hours in the year (or in a season) rather than first picking the worst hour of the day. (In addition to a bit more "drying"), reasons for this suggestion are to reduce the tendency of this metric to be dominated by the first few hours after sunrise. Generally I think a day where its hazy all day is more objectionable than one where its hazy at sunrise (which probably

makes it "prettier"). A move away from the single worst hour would also help screen out effects of instrumental measurement errors, for an indicator which requires combining information from 3 separate instruments.

- Note that many of the above suggested revisions or sensitivity analyses all tend to essentially cut off many of the highest humidity and poorest visibility hours, and so they might logically be combined with a higher percentile form (like the 98th).
- Consider moving quickly to initiate a small pilot urban visibility monitoring network to identify and work out the bugs in the proposed methods. While there has been considerable experience in the visibility research community with the operation of ambient nephelometers and aethalometers – in some cases equipped with switching heads that alternatively sample different particle sizes – such combinations have not really been tried in routine network operations, and it would be critical to get some early experience in this area.

September 2009 PM Visibility Risk Assessment, Charge Questions:

1) After careful consideration of the evidence provided in the second draft ISA, and in particular the significant body of work that has been conducted by the Regional Planning Organizations under the Regional Haze Rule, (i.e. information on urban and rural PM concentrations and compositions), we have decided to continue to focus this assessment on the PM induced visibility impairment that is occurring in urban areas. What is the Panel's view on this approach? Is the rationale supporting the selected approach clear and appropriate?

Yes, focusing a secondary PM NAAQS in a way that is complementary to the existing Regional Haze Rule (RHR, which applies in Class I National Parks and Wilderness Areas), is appropriate, efficient and the rationale is logical and clearly stated. A somewhat semantic point here is that the physical areas occupied by "Class I" Federal lands and by "urban" areas each cover relatively small fractions of US land area, and that a secondary standard "complementary to the RHR" should apply to all "non-Class 1 areas", for which "urban visibility" is a effective descriptive term that should be intended to represent larger geographical areas and larger ranges of urban, suburban and rural non-class 1 area population densities. In addition, there is no need for the RHR and secondary (or primary) PM NAAQS to be viewed as mutually exclusive regulatory mechanisms. Steps taken to meet a secondary (or primary) PM (or SOx or NOx) NAAQS applied to "urban" areas may well result in improvements in regional haze in remote class 1 areas – and vice versa.

2) After further considering the nature of urban versus more remote area PM, and in light of discussions with CASAC at the April 2, 2009 meeting, we have decided not to develop an urban optimized algorithm at this time, but instead to rely on the original IMPROVE algorithm to relate urban PM to local haze (PM light extinction). Is the Panel generally supportive of this approach? Is the rationale supporting this decision appropriate and clearly presented?

Yes, this is a very reasonable approach at the present time. Its likely that improvements may be made to this equation in the future – especially as the new urban extinction data become

available, but its highly unlikely that alternative equations would result in major revisions to results obtained from the IMPROVE equation. As a practical matter, the equation becomes irrelevant (except for purposes of source attribution) as the standard is implemented, since the optical indicator essentially aggregates the cumulative effects of all the contributing pollutant species without requiring any information on the individual species concentrations or effects.

> *3) In a change from the planned approach presented in the Scope and Methods Plan, we have decided to conduct a reanalysis of the urban visibility preference studies available at the time of the 2006 PM NAAQS review, rather than conducting new public preference studies since it is highly unlikely that the results of new studies could be completed in time to inform this review. This reanalysis was designed to explore the similarities and differences (comparability) between the current studies and to assess what information could be drawn from these results to inform the selection of visual air quality (VAQ) candidate protection levels (CPLs) to be used in subsequent impact assessments. This reanalysis also includes a recent study by Smith and Howell (2009) for Washington, D.C. which was presented to the CASAC during the public comment phase of the April 2, 2009 meeting and later provided to EPA staff. Does the Panel agree that the information provided by this reanalysis is useful to inform the selection of CPLs? Does the Panel agree that inclusion of the Smith and Howell (2009) is appropriate in both the ISA and Visibility Assessment? To what extent does the Panel consider that the reanalysis of the urban visibility preference studies is clearly and appropriately characterized?*

While the Agency is encouraged to allocate sufficient funds and resources to support the planned new urban visibility preference studies, the decision to conduct a careful reanalysis of currently available urban visibility preference studies – including the recent Smith and Howell (2009) data - was the most reasonable approach to developing a range of candidate protection levels (CPLs) for use in the current risk assessment. There would not have been time to complete the new studies quickly enough to support the current assessment, and the results of those new studies will be more useful if they are carefully planned and executed, and not rushed according to an accelerated NAAQS review schedule. The reanalysis shows a relatively strong degree of convergence in the identification of unacceptable levels of visual air quality across the different study areas, and provides a sufficient basis for use of the proposed CPLs in the current risk assessment, and for considering alternative levels of secondary PM standards based on PM light extinction.

> *4) We have chosen to use the range that represents the 50th acceptability criteria across the four cites studied (i.e., the VAQ level that best divides the photographs shown into two groups: those with a VAQ rated as acceptable by the majority of the participants, and those rated not acceptable by the majority of participants) as CPLs to characterize the nature of the impact on urban VAQ associated with current PM levels. Please comment on the clarity and appropriateness of the rationale supporting this decision. Does the Panel have suggestions for alternative ranges to consider?*

Use of a range representing the 50th percent acceptability criteria for VAQ levels across the four study areas is an appropriate approach, for which the justification is logical and clearly stated. It is likely that these “acceptable” and “unacceptable” VAQ levels will be further refined in the future, as results of additional studies become available. But for use in the current NAAQS review, the

identified range of VAQ levels provides a sound and defensible basis for establishment of a separate secondary standard based on achieving and maintaining acceptable levels of VAQ in urban areas.

5) A new indicator relating ambient PM to urban VAQ (i.e., PM light extinction) has been evaluated to improve our characterization of the relationship between ambient PM and visibility impairment. Is the Panel generally supportive of this approach? To what extent have we provided an adequate justification for the new indicator used?

Yes, the proposed new indicator is logically justified and clearly described. The PM light extinction indicator will be directly and instantaneously responsive to changes in the concentrations of the PM components which affect it, and would have the eloquent advantage over all other currently or previously considered NAAQS indicators in that the indicator, in this case would be the effect.

6) An averaging time of one hour as a practical minimum time period was used in this assessment in recognition that, while the visibility impacts are nearly instantaneous, the urban VAQ does not generally change significantly from minute to minute, but does vary from hour to hour. To what extent does the Panel support this approach? Does the Panel consider the rationale supporting this approach to be clearly and appropriately presented?

I have no major objections to the use of a one-hour averaging time, and think logical arguments for it have been clearly presented here. I don't completely agree, however, that since visibility effects are perceived instantaneously, a one-hour averaging time is necessarily the best or most effective averaging time <u>for a regulatory metric</u>. Use of somewhat longer averaging times – such as the 4-hour intervals, (selected after careful consideration in both Phoenix and Denver) would tend to help screen out effects of short-term outliers, unusual weather, instrumental noise or malfunction, and would also tend to place more emphasis on causes of poor visibility which are more persistent than ephemeral.

I'm also not sure I agree that the use of the single worst (daylight) hour in a day (prior to calculating a percentile) is necessarily better than applying that same (or another) percentile to all the hours of the year (or of the season). Is the 95th percentile of the worst hour of the day a better metric than the 98th percentile of of the second worst hour, etc.? Time permitting, it might be informative to conduct a sensitivity analysis, exploring effects of some of these alternative combinations.

7) We have chosen to use the 90$_{th}$ and 95$_{th}$ percentile forms in our assessment of alternative secondary (welfare-based) standards. Please comment on the use of these alternative forms.

The use here of the 90th and 95th percentiles is reasonable for purposes of this first draft assessment, and helps illustrate the important effects that selection of the form can have on the frequency and magnitude by which certain levels are exceeded (or not). That being said, I don't think either of these (or any other percentile) can be uniquely well-justified compared to other forms. In the last PM NAAQS review, EPA staff had initially suggested the 90th percentile, as

being "consistent" with the regional haze Rule (which focuses on the worst 20 % of days, for which the average is roughly equal to the 92nd percentile. However, the flaw in this "logic" is that the RHR specifically focuses on and requires improvements in these worst days, while using a similar percentile as the form for an urban visibility standard would essentially discard these worst days as unimportant. The 90th percentile also seems like an illogical metric to combine with the worst hour of the day (the whole day is unacceptable if even a single hour exceeds the threshold, but this only becomes unacceptable if it happens more than 37 days a year).

Off hand, I don't think its necessarily logical that a welfare standard should be set at a lower (less stringent) percentile than a ("more important") health standard. By this logic, secondary standards should always be set more leniently (or equal to) the "more important" primary standards (as they, ah, have been…). If adverse effects on either health or welfare are experienced during single or a few high short-term exposures, then the percentile forms of both kinds of standards should be set accordingly. From that perspective, the 98th percentile might also be considered as appropriate for protecting visibility, and should not be considered unreasonable without better justification. A (much) more leisurely pace toward attaining a "less important" secondary standard is accommodated by less urgent implementation requirements.

At the same time, since there is no compelling basis for selecting a specific (relatively high) percentile, this might currently be considered as an area where there is "flexibility" to balance alternative percentiles with other elements of the standard to find the best combinations of indicator, averaging time level and form which would lead to maximum improvements over different geographical regions but which would also be feasible to attain over reasonable time periods (or accompanied by guidance outlining schedules of reasonable progress in the implementation phase).

> *8) To what extent does the Panel support the graphical displays presented in this chapter? As currently presented, do these figures clearly summarize the assessment results? We have combined data from multiple studies for two locations - British Columbia and Washington, D.C. - (Abt Associates, 2001; Smith and Howell, 2009 test 1) as presented in Figure 2-14. Does the Panel agree with developing a composite dataset for each of these two urban areas?*

The graphic displays (such as those in Figures 2-2, 2-3, 2-5, 2-7, and 2-9 through 2-14) which compare visibility levels in deciviews with percent respondents rating the VAQs acceptable are an excellent way of showing results from the different studies in common units, showing the range or distributions among different respondents and across the different study areas, and for showing the convergence of results for acceptable/unacceptable visibility in the range of 20 to 30 dv. The use of composite data sets based on combined study results for the British Columbia studies and for the Washington DC studies is also a logical approach, and seems appropriate for combining these study areas with results from the Phoenix and Denver studies.

> *9) Despite significant differences in study characteristics (e.g., size, location), to what extent does the Panel support combining and comparing the results from the four cities, as shown in Figure 2-14? What is the Panel's view on the clarity and adequacy of the descriptions of the uncertainties and limitations associated with such a combined*

assessment and the conclusions that can be drawn from the assessments? Please provide comments on additional insights, uncertainties, or caveats that should be considered.

I think combining results from the different study areas is an appropriate approach that takes maximum advantage of all the currently available data. To a certain extent, the relative convergence of results from the different study approaches and study areas strongly supports the use of these data collectively for use in identifying a relatively narrow range of acceptable VAQs across many locations and respondents. Although a formal uncertainty analysis was not conducted (nor do I believe it would have been very informative here), I think the descriptive information on uncertainties and limitations of this approach conveys that information clearly. It should also be noted that the uncertainties associated with this approach for selecting a level or range of acceptable VAQ are likely to be relatively small compared to uncertainties associated with other aspects of a secondary PM standard such as the averaging time, representing a day by its worst 1 (or more) VAQ hour(s), and the frequencies (percentiles) for which unacceptable VAQ conditions are acceptable.

While the focus here has been on discerning a decision point or range of VAQ conditions considered acceptable vs. unacceptable, it should also be recognized that there are welfare benefits associated with improvements across all parts of the visibility spectrum, including for example a shift of days with "good" to "excellent" VAQ. As a practical matter, efforts made to improve visibility conditions in any part of the distribution are likely to shift the entire distribution, with benefits that will be understated by a single 50% acceptability approach.

10) We have used the combined results presented in this chapter to develop a range of CPLs that are used in subsequent steps of the assessment. To what extent does the Panel support the range of CPLs used and the justification provided for selecting this range? Does the Panel recommend consideration of any alternative approaches or criteria for selecting CPLs?

As indicated above, the chapter presents a clear justification for the range of CPLs used in subsequent steps of this assessment. Its likely that results of future urban visibility preference studies can help refine (and may well lower) this range in the future, but the proposed range is logically derived from currently available research results and appropriate for use in evaluating and proposing a range of PM light extinction levels that could form the basis of a secondary standard to protect urban visibility in the current NAAQS review cycle. While other approaches for selecting alternative CPLs might be considered, I think the current range is reasonable, and that a lot of additional work in this area is not currently justified, considering the relatively large influence that other elements of a standard (such as averaging time, percentile form, RH screens, etc.) will have on the ultimate effectiveness of the standard.

11) Overall, we consider this assessment useful for providing information for the design of future urban visibility preference studies. Does the Panel support this conclusion and does the Panel have any recommendations for changes that could be made in the discussions of this information to enhance its usefulness for this purpose?

Yes, I would agree that this assessment provides a very useful basis for identifying designs for

and priority information needs from future urban visibility studies. As indicated above (and supported by the discussions of issues in this assessment), such studies should be designed to provide information that will help refine other elements of a PM light extinction standard in addition to identifying the level(s) of VAQ considered unacceptable across a range of different urban locations and viewing conditions. For example: what frequency of poor visibility conditions is acceptable; should this frequency be applied to the single worst hour of a day or should days with prolonged periods of poor visibility be considered more adverse than days with poor visibility for only an hour or two; what are preferences for shifts in other percentiles of the visibility frequency distribution – such as an increase in the frequency of very clear days, etc?

As always, there is a need to carefully balance the advocacy of future research needs with the confidence in using the best currently information to make sound decisions to better protect human health, welfare and environment. In this case, currently available information is more than adequate to establish a basic secondary standard based on PM light extinction, even while future research is clearly needed to refine such a standard in subsequent NAAQS review cycles.

It should also be noted that the more detailed information on hourly light scattering and absorption that will result from implementing a standard with a PM extinction indicator, will itself provide an invaluable information resource that will help support future urban visibility preference studies and refine light extinction equations for different mixes of urban aerosols. These data would also be available for near-real time reporting and public communication, would be useful for air quality forecasts and (in combination with other continuous measurements) will provide added information on highly time-resolved fine & coarse particle composition, including black carbon, of value for various health effects studies as well).

> *12) Are the goals articulated in the first paragraph of this chapter achieved in the remainder of the assessment? If not, does the Panel have suggestions for additional assessments that should be done?*

Yes, this represents a very thorough and clearly presented analysis, and is extremely informative for a first draft. As indicated elsewhere, I think the RH limit should be used as a screen (discard hours above it) rather than a cap (set hours above it equal to it) and should be set at 90% rather than 95%. Other possible metrics that might be explored – if time and resources permit - in a second draft assessment might include: use of 2nd highest daylight hour, or the max 2 to 4-hour average (combined with higher percentiles), calculating percentiles based on all daylight hours in a year or a season, shortening the daylight window to exclude the (typically most humid) hour(s) just after sunrise and just before sunset, etc.

> *13) Are the methods and approaches taken in these assessments, including those for monitor site selection, incomplete data adjustments, and the use of the CMAQ model to augment speciation data, appropriate and is the rationale for their selection clearly articulated?*

All of the above are reasonable and were clearly explained. Results from a few of the selected cities don't seem quite right to me, but this tends to make me suspicious of the data, rather than the methods, as the latter seem logical to me. For example, the coarse mass from St. Louis

doesn't make sense, and I'm surprised by (& suspicious of) the poor correlations between 4-hr afternoon B_{ext} and fine mass in Table D-1 and Figure D-2 for St. Louis ($R^2 = 0.27$) and Philadelphia ($R^2 = 0.36$) – although I'm sure these will improve if an RH screen (not a cap) were applied to the B_{ext} estimates.

14) Is the approach used to estimate PRB as described in chapter 3 and Appendix C appropriate?

The PBR calculations look fine, and I've noticed that for most species they seem to agree reasonably well with the carefully derived estimates of "natural background" at nearby IMPROVE sites for the Regional Haze Rule. For other species (nitrate and fine soil) the PBR estimates seem (illogically) much lower than the IMPROVE natural background estimates. I don't think this is actually very important, but it might be informative to compare these (small numbers in both cases) if time allows.

15) We consider the results generated by these analyses to be reasonable based on PM composition and relative humidity data. Does the Panel agree? Are there other tests of reasonableness that could be applied?

Generally they look reasonable to me. Possibly you could compare estimated hourly extinction estimates for Phoenix with the transmissometer data there. For Atlanta, the hourly species estimates might be compared with the hourly SEARCH data there if the time periods overlap. Similar comparisons might be made with the hourly species composition from the Supersite data from Fresno, St. Louis and Pittsburgh. For many locations, the airport ASOS data are actually not bad, especially if the raw data can be accessed prior to binning and truncation. Even with the censored data, some comparison might be made for hours when very poor visibility (< 10 km) was reported (or predicted), again with RH screening. See also #s 16 & 17 below.

16) In addition to a qualitative discussion of possible sources of uncertainty and variability, are there quantitative methods for addressing uncertainty and variability associated with these assessments that the Panel would recommend?

The qualitative discussions are helpful. I think some quantitative comparisons between measurements and measurements such as those suggested in #15 above could help convey some of the uncertainties. I consider the comparisons between 4-hr afternoon B_{ext} and fine mass in Table D-1 and Figure D-2 as good illustrations of causality, the similarities and differences and variability between a dry and wet (ambient) indicator – during the generally drier daylight hours, and also as a form of QA that (for me) adds confidence in the estimates. If the B_{ext} in those comparisons were further dried by use of 90% RH screen, the fit and our resulting confidence should both improve.

As indicated above, I think the graphic displays (such as those in Figures 2-2, 2-3, 2-5, 2-7, and 2-9 through 2-14) which compare visibility levels in deciviews with percent respondents rating the VAQs acceptable are an excellent way of showing the variability in results among different respondents and across the different study areas, as well as showing the convergence of results for acceptable/unacceptable visibility in the range of 20 to 30 dv. This provides a quantitative

sense of the range of the uncertainty that selection of any single VAQ level will have in representing responses of all viewers in all areas.

Elsewhere in these comments I suggest alternative metrics for some of the parameters used here. If time allows, quantitative comparisons of results from these "sensitivity runs" might also give some indication of variability and uncertainties associated with selecting any approach among several that are approximately equally justifiable.

> *17) A number of appendices are provided at the end of this document. Does the Panel agree that this information is useful to retain? Does the Panel agree with the level of detail provided in the body of the report and its organization and distribution throughout the document.*

I find the information provided in the appendices very informative – it really helps to "see" and gain confidence in the complex data and relationships which underlie the resulting estimates. I think the level of detail, the organization of information and its distribution within the assessment document and appendices is excellent. This is very well conceived, organized, executed, and written – a pleasure to read!

As indicated in previous comments, I'm suspicious of some of the poor correlations in Appendix D and hopeful that the recommended 90% RH screen will improve those results. Also, It would help if some justification were provided for the use of the LOESS regression in Appendix D. I note that most of the regression lines are nearly straight, while a few curve up and others curve down, but I'm not aware of the theoretical justifications for these differing non-linear relationships, and would also like to have a better sense of how a LOESS R^2 of 0.8 compares to that from an ordinary least squares approach. Also if you used a linear regression, you could also report the slopes and intercepts (if any) in Table D-1, which would convey added information of interest.

> *18) Does the Panel agree with the approaches used to simulate just meeting air quality conditions for the current and alternative PM standards? In particular, is use of the proportional rollback approach appropriate in the context of the urban PM visibility assessment?*

I think this approach is reasonable and have no alternatives to suggest. In reality I don't think we would expect all species to be rolled back by the same proportion, but don't think there's much basis to assume otherwise. Table 4-7 is an especially informative display of relevant information! I assume that for the final policy assessment it would be possible to add results for other PM combos like 13/30, etc.

You show the (90th and 95th percentile) B_{ext} levels associated with just meeting various $PM_{2.5}$ standards. I wonder if (at least for one or two urban areas) it might be possible to provide an indication of what the distributions of hourly B_{ex} levels would be under current conditions and associated with just meeting the alternate PM standards – (and might these also be compared with the distributions associated with just meeting the alternative B_{ex} standards)? Then, if there are Winhaze (or sufficient actual haze cam photos) images available, maybe some percentiles

along these distributions be graphically illustrated. This might be an effective way of communicating the additional benefits that the optical standard might provide – especially at the clearest and haziest ends of the distribution. Unfortunately, the available Phoenix and Dallas WinHaze photos would likely be poor choices for such illustrations since they are relatively clean to start with, although you could increase extinction there up to the levels allowed by the current standards.

> *19) To what extent does the Panel consider the presentation of "what if" scenarios for retention of the current secondary PM2.5 NAAQS and consideration of alternative, more protective secondary NAAQS in sections 4.2 and 4.3 to be clearly written with an appropriate level of detail? Do the correlation analyses presented in Appendix D provide sufficient insight into the suitability of alternative indicators based on sub-24 hour averaging periods for PM2.5? Are there additional alternative standard scenarios that should be evaluated?*

As indicated above, I suggest using a 90% RH screen and several other "sensitivity runs" that will generally tend to dry out the aerosol, somewhat reduce the current East/West differences and result in lower extinction estimates. It would be informative to see the results of some of these "softer metrics" compared to the alternative primary PM standards. Since those changes would essentially screen out many of the worst visibility hours from a regulatory metric, I think it would also be appropriate to consider combining them with a higher percentile form – such as 98%tile, which I think can be as logically justified (or more so) as (than) the 90th or 95th.

A related point - a bit off-topic here, but maybe appropriate for the policy assessment (which I haven't read yet) - is that since secondary standards have no required time fuse, it should not necessarily viewed problematic if some (eastern) areas have much higher "residual secondary non-attainment" (hours above the B_{ext} standard after primary standards are attained) than other (western) areas, as this would reflect actual visibility conditions and EPA guidance might suggest the concept of a staggered time schedule where dirtier areas have more time and vice versa such that all areas might be expected to show similar "rates of progress", as in the Regional Haze rule. Such a "progress-based" secondary standard would allow better coordination with the Regional Haze Rule, and would better accommodate different visibility, pollutant, humidity and/or scenic charistics in different regions.

Specific Comments on September 2009 PM Visibility Risk Assessment

p. 1-7, lines 16-20: I don't fully agree with the "logic" that since visibility is worse when RH is highest, then greater protection is needed during the times when humidity is highest.

p. 1-10, line 28: Add "s" to "characterization" or change "were" to "was".

p. 2-14, lines 16-18: Might it be possible to consider compliance metrics which – like in Phoenix – are not based on a single absolute threshold, but rather based on a required shift in the distribution. Arguably, there are important welfare benefits in shifting from moderate to good and from good to excellent visibility, this kind of approach would be more consistent with the Regional Haze Rule, and it would be more accommodating of regional differences in particle composition and humidity.

p. 2-26, lines 17-36: Considering the extensive preceding assessments to derive an appropriate range of levels for a visibility standard, only a short paragraph is devoted to identifying a proposed averaging time and form. I don't think either the 1-hour average or the 90 or 95th percentiles are uniquely well justified here, and in fact seem somewhat inconsistent with each other ("its an intolerable day if visibility is impaired even for an hour, but I don't mind until it happens more than 37 days a year"). One effect of the 1-hour averaging time (considered over all daylight hours) is to assure that the day's maximum (or near max) humidity is encountered, often also at the period of lowest wind speed, lowest mixing height – and often at a period when its most difficult to discern between morning fog and pollution effects. In many mountain/valley settings, winter visibility may be poor for only an hour or 2 just after sunrise, but much clearer throughout the rest of the day. But unlike a short-term health effect, one hour of bad visibility does not necessarily ruin your whole day, and a day when visibility is impaired all day long should logically be considered more adverse than a day with one bad hour. So I see no logic in calling the entire day impaired if it has one bad hour, and a percentile form – like 90th, 95th or 98th – might more logically be applied to all the daylight hours, rather than to the single maximum hour.

It might also be noted that both the Denver and Phoenix standards and what EPA staff proposed last time are all based on 4-hour averaging times. Yes, visibility is perceived instantaneously, but arguably impairment that persists over longer periods is more objectionable than short-term aberrations, when it may be most difficult to discern the difference between pollution and natural influences. A somewhat longer averaging time would also help minimize effects of measurement errors, and would have the effect of reducing, somewhat, the influence of RH on the regulatory metric.

p. 3-8, Figure 3-3: You could add the dates (2000-2004) to the figure caption.

P 3-8, line 12: Add "and high relative humidity" to this list. You might also add "This is also the time of day when ground fog is most prevalent, and when its most difficult to discern the difference between natural and manmade causes of impairment."

p. 3-10, Table 3-2: Its seems curious that for all cities where the site-specific value is not the same as the design value, the site-specific values for both annual and 24-hour are always greater than the site-specific.

p. 3-16, line 14: I assume that the "old" IMPROVE conversion factor of 1.4 x OC is used here, right? Was any adjustment used to account for differences between IMPROVE and CSN OC and EC data? If so, specify; if not, some explanation is warranted.

p. 3-17, Figure 3-4: The time intervals (X-scales) are incorrect. Also, I'm surprised to see the winter Detroit EC & OC are so much higher for evening rush hour than for morning rush hour. I don't think that's a typical pattern – or at least would be different in areas where winter inversions are common.

p. 3-21, lines 1-10: It would be interesting to see how the SANDWICH organic matter compares to 1.4 (or 1.8) x OC.

p. 3-25, lines 22-28: Now I'm confused – I thought you were going to exclude hours with RH>95%. Now it sounds like you will "cap" them by setting the f(RH) for hours > 95% to the f(RH) at 95% - which is something like 7.4. I don't like this at all, as one of the good reasons to exclude such hours

is to exclude periods of fog or precipitation. Its like you're saying, lets pretend its not raining but that its just really humid. As indicated elsewhere, I think it would have been useful to conduct this exercise with alternative constraining metrics – such as a 70% (Denver), 80%, and/or 90% (Phoenix) RH cutoff, a 4-hour average, a shorter and consistent (8 AM to 4 PM) daytime window, and/or percentiles based on all the daylight hours rather than just the worst hour in a day.

p. 3-28, Figure 3-7b: The $PM_{10-2.5}$ data for St. Louis don't seem like they could possibly be very representative. Also, I'm reminded of seeing some of the coarse particle composition data from Bob Vanderpool and noting the very high coarse EC content in cities with steel mills like Birmingham – where using the IMPROVE coarse scattering efficiency may understate effects.

p. 3-30: It looks to me like a metric based on all the daylight hours would be more stable and less variable across the different sites than the daily max hour.

p. 3-35: This figure helps illustrate why a RH cutoff (not a cap) at 90% RH (or lower) would result in a temporally and spatially more stable metric than capping at 95%. I wonder how this figure would look if you limited it to only the daily 1-hr max B_{ext} values?

p. 3-37: Was an hourly f(RH) used in the PRB estimates?

p. 4-1, lines 11-12: I disagree with the recommendation that (only) a PM_{10} inlet be used. Neither the nephelometer or aethelometer respond as efficiently to coarse particles as to fine ones, and the measurements with a 10 micron head would not truly represent the total PM scattering or absorption. At a minimum, I would consider a "switching" inlet which alternately samples at PM_{10} and $PM_{2.5}$ (or PM_1). If that doesn't work, I would use a 2.5 or 1 micron inlet only and rename the indicator "fine particle light extinction". Coarse particles contribute much less in most urban areas, Phoenix already has a visibility standard, and arguably coarse particles could be excluded with the argument that concentrations are much less spatially uniform than fines and so a point measurement is unlikely to be representative of areal conditions (see your St. Louis data to illustrate this point). A possible alternative would be to "allow" use of fine cut neph +aeth at locations where PM coarse extinction is estimated to be less than 10% on hazy days and require transmissometers elsewhere.

An additional question here is that since it may take some time to implement a new network to measure PM light extinction, are there any ways to estimate PM light extinction from existing measurements? I would think reasonable estimates could be derived from a combination of continuous PM_{10}, $PM_{2.5}$, RH and a "generic aerosol f(RH)" growth function. This f(RH) function might be reduced or eliminated for periods when it could be otherwise demonstrates that non-hygroscopic aerosols, such as wood smoke or soil dust were prevalent. Another alternative approach would be for EPA to work collaboratively with NOAA to improve the quality and accessibility of airport ASOS visibility data. This would reduce costs for a new "urban" visibility measurement program, as there are many hundreds of such sensors in operation throughout the country, and since the B_{ext} threshold for the ASOS sensors (in its original form before it gets censored and binned) is about 50 Mm^{-1}, well below even the cleanest level of 74 Mm^{-1} (20 dv), currently being considered as a level for a secondary PM NAAQS. This approach would have the added benefit of improving aviation safety.

p. 4-5, Table 4-3: I wonder if it would be possible to calculate the 24-hr & annual $PM_{2.5}$ concentrations & changes that would be associated with meeting these different optical limits.

p. 4-7, Table 4-4: Would it be possible to calculate and show the B_{ext} levels & changes that would be associated with these PM limits?

p. C2, Table C-1: It might be informative to compare the IMPROVE "natural background" calculations to these modeled estimates. For some species (SO4, OC, EC) I think the estimates are similar but for nitrate and soil, the IMPROVE natural background levels (which should be lower) are substantially higher.