Table I:
Dissecting EPA's Response to the Problems of Self-Administration and Self-Reporting of Respiratory Function Tests

EPA Response [USEPA (2008c, p. 33)	*Non Sequiturs and Errors*
Finally, NAM contends that EPA did not recognize the fundamental data quality problems with self-reported respiratory testing found by Kamps et al. (2001).	We claimed that both self-reporting and self-administration created serious information quality problems. EPA's description captures self-reporting but ignores self-administration.
In section 7.2.3 of the Criteria Document, EPA does in fact state that PEF measurements have been shown to be more variable than FEV1 in some studies (Vaughan et al., 1989; Cross and Nelson, 1991) and can have an element of uncertain reliability when self-administered by study subjects.	*Non sequitur #1*: Variability across measurement instruments is a well-known phenomenon, and we did not allege that its existence constituted an information quality defect. *Error #1*: Vaughan et al. (1989) did not rely on self-administration. *Error #2*: Cross and Nelson (1991) concerns asthma diagnosis and treatment, not self-administration or self-reporting.
However, Lippmann and Spektor (1998) state that PEF measurements from small inexpensive flow meters, which are more feasible to use in field studies, can produce similar results to PEF measured spirometrically.	*Non sequitur #2*: Lippmann and Spektor (1998) compared mini Wright peak flow (mWPF) readings with spirometric peak expiratory flow rates (PEFR). They did not address either of the information quality problems we alleged.

Table II:
Severity of Any Spirometric Abnormality Based on FEV_1

Degree of severity	Percentage of Predicted FEV_1
Mild	> 70
Moderate	60-69
Moderately severe	50-59
Severe	35-49
Very severe	< 35
Source: Pellegrino et al. (2005, p. 957, Table 6).	

How should these severity scores be used? The ATS guidelines explain:

> Severity scores are most appropriately derived from studies that relate pulmonary function test values to independent indices of

> performance, such as ability to work and function in daily life, morbidity and prognosis. In general, the ability to work and function in daily life is related to pulmonary function, and pulmonary function is used to rate impairment in several published systems. Pulmonary function level is also associated with morbidity, and the patients with lower function have more respiratory complaints.[79]

This is not how epidemiologists have used lung function test data. They do not use these data to assign subjects into a handful of categories. They use the data to divine vanishingly small group decrements.

The ATS interpretative guidelines also speak to the question of how small changes should be interpreted. The clear theme is caution. For normal subjects, a change in FEV_1 less than 5% within a day is not likely to be significant.[80] Looking at the relevant controlled human subjects studies, Adams (2006) reported changes greater than 5% for 80 ppb ozone at 5.5 hours, but not at 60 ppb or 40 ppb. Schelegle et al (2009) also reported changes greater than 5% at 80 ppb, but not at 70 ppb or 60 ppb.[81]

In Administrator Johnson's 2008 decision, considerable weight was given to group mean difference in FEV_1 of 2.6% per 50 ppb ozone; by interpolation, this is approximately 0.5% for the difference between 84 ppb (the practical meaning of the 1997 standard) and 60 ppb (the lowest value under consideration for the 2008 revision). Differences this small are hard to credit as meaningful effect sizes. EPA has assumed that if a phenomenon can be measured, it must be important. By implication, EPA has concluded that effects too small to be clinically meaningful for an individual are nevertheless environmentally crucial in a population.

C. Validity and reliability problems arising because of potential investigator bias.

In the 2008 RFR, we discussed the matter of how research subjects were "coached" in the performance of lung function tests. Because diagnosis and assignment into perhaps a half dozen categories is the clinical purpose of these tests, coaching is specifically required by the ATS technical guidelines. We might presume that subjects in controlled experiments were coached in similar ways to avoid investigator bias, but we do not know for sure. As for observational epidemiology studies, our knowledge about investigators' coaching practices is completely blank because they did not disclose this information. Variations in coaching

---

[79] Pellegrino et al. (2005, p. 957, internal citations omitted).

[80] Pellegrino et al. (2005, p. 961). For COPD patients, the intraday change must exceed 11%.

[81] The ATS guidelines define as significant week-to-week changes greater than 12% in normal subjects. They are silent concerning day-to-day changes, which are the subject of the observational environmental epidemiology studies.

can be expected to have a material effect on the data; indeed, the ATS guidelines say it does.[82] Variations in performance are much greater for self-administered tests.

D. Validity and reliability of self-administered lung function tests

Because EPA's Response to Comments merely recycled text from the 2006 Criteria Document, we have taken a closer look at that text in context. The Criteria Document contains an extensive discussion of data from pulmonary testing, and the many studies that rely on it for estimating the effects of air pollutants.[83] But the only place in these 250 pages where there is even a suggestion that self-administered testing is problematic is in the snippet of text that EPA reproduced in its Response to Comments. In the Staff Paper, even this tidbit is absent.

We believe an appropriate additional remedy that EPA can complete now is to examine each article referenced in the 2006 Criteria Document that relied on self-administration of respiratory function tests, and answer the following questions:

- Did the researchers report that subjects were trained in the use of the lung function test technology?
- Did the researchers report having validated that this training was successful?
- Did the researchers report having validated the data that subjects provided?

Affirmative responses to these questions would increase the trustworthiness of study results. Negative responses, however, would strongly suggest that it is inappropriate to take the results at face value, as EPA has done. [84]

E. Validity and reliability of critical self-reported data, including lung function test data

EPA's Response to Comment acknowledged that self-reporting created information quality problems. As Table I makes clear, EPA's response was a combination of non sequitur and incorrect statements, which we pointed out in the 2008 RFR.

---

[82] Miller, Hankinson et al. (2005, p. 323): "Throughout the manoeuvre, enthusiastic coaching of the subject using appropriate body language and phrases, such as 'keep going', is required." Children especially need to be effectively coached, requiring special skills and a child-friendly environment. *See also* Miller, Crapo et al. (2005, p. 158): "Perhaps the most important component in successful pulmonary function testing is a well-motivated, enthusiastic technician."

[83] U.S. Environmental Protection Agency (2006a, Sections 6 [54 pp] and 57 [195 pp]).

[84] Questions such as these should be part of the pre-dissemination review that OMB's government-wide information quality guidelines require and which EPA promised to perform. That these questions appear never to have been asked before, much less answered, suggests that EPA's actual pre-dissemination review program is nonexistent.

We believe an appropriate additional remedy that EPA can easily complete now is to examine each article referenced in the 2006 Criteria Document that relied on self-reporting, and answer the following questions:

- Did the researchers report that subjects were trained in the accurate reporting of respiratory function tests and other data?
- Did the researchers report having validated that this training was successful?
- Did the researchers report having verified that subjects recorded data correctly and contemporaneously?
- Did the researchers report having validated the data?

As before, affirmative responses to these questions would increase the trustworthiness of study results, and negative responses would raise red flags.

F. Discarded inter-maneuver variability

Vaughan et al. (1989), compared FEV_1 and PEFR measurements across different test instruments to ascertain their relative merits for the clinical purpose of diagnosing pulmonary impairment.[85] The authors reported the standard deviations across maneuvers for the lung function tests themselves (FEV_1 and PEFR for the Jones Pulmonor Spirometer: FEV_1 for the mini-Wright peak flow meter). We reproduce these data in Table III.

The magnitude of these standard deviations is similar to the effect sizes that laboratory researchers and epidemiologists have been reporting due to ambient ozone levels below 75 ppb. For example, the highest average FEV_1 decrements Schelegle and colleagues report at 70 ppb and 60 ppb are about 5% and 2%, respectively. They characterize the decrement at 70 ppb as statistically significant, but this result depends on the assumption that each FEV_1 measurement is fixed and has zero variance. Similarly, Mortimer et al. (2002) report fractions of a percent change in PEFR in children that appear to be statistically significant, but this too depends on the assumption that each PEFR measurement for each child is fixed and has zero variance.[86]

We believe an appropriate additional remedy that EPA can complete now is to examine each article referenced in the 2006 Criteria Document that relied on a lung function test, and answer the following questions:

---

[85] This study was mentioned in the 2006 Criteria Document. EPA raised it again in its Response to Comments in response to an unrelated issue, which prompted a closer look.

[86] Mortimer et al. report that "[t]he maximum of three manoeuvres, performed while standing, was recorded." They excluded values below 70 $L\text{-}min^{-1}$ and above 450 $L\text{-}min^{-1}$, so the magnitude of inter-maneuver variance could be substantial. Respiratory function tests were performed by the children themselves, who Mortimer et al. report "were trained."

- Did the researchers report having followed the ATS guidelines in the conduct of lung function testing?[87]
- What quantitative criterion did the researchers use for determining whether a maneuver was "acceptable"? What percentage of the FEV_1, PEFR, or other lung function measurement is this?
- Did the researchers incorporate inter-maneuver variance in their statistical analysis, or did they discard it in favor of the maximum or a central tendency measure such as the mean?
- Did the researchers report the FEV_1, PEFR, or other lung function measurement for each maneuver?[88]
- ATS guidelines call for retaining the results of each "acceptable" maneuver.
- For each study in which the researchers claim to have followed ATS guidelines, EPA should formally request that the researchers publicly disclose this information, with appropriate censoring of identities to ensure that privacy is protected.
    - For each study that was EPA-funded, EPA should formally instruct the researchers to disclose this information, as provided for by OMB Circular A-110.[89]
    - For each study that was funded by a different federal agency, EPA should formally ask that agency to issue an instruction to disclose this information, as provided for by OMB Circular A-110.

---

[87] *See* Miller, Hankinson et al. (2005, p. 325).

[88] With this information, any qualified third party (including EPA) can perform the analysis again to determine whether inter-maneuver variation has a material effect on standard errors and statistical significance. This kind of analysis can be done easily, and is explicitly called for in EPA's Information Quality Guidelines, which

> provide for the use of especially rigorous "robustness checks" and documentation of what checks were undertaken. These steps, along with transparency about the sources of data used, various assumptions employed, analytic methods applied, and statistical procedures employed should assure that analytic results are "capable of being substantially reproduced."

See U.S. Environmental Protection Agency (2002, p. 47). An inability to substantially reproduce is an information quality defect per se. The absence of actual robustness checks is strong evidence of negligent pre-dissemination review.

[89] Office of Management and Budget (1999, __ 36(c)): "The Federal Government has the right to (1) obtain, reproduce, publish or otherwise use the data first produced under an award; and (2) authorize others to receive, reproduce, publish, or otherwise use such data for Federal purposes."

With this information, EPA can perform the robustness checks required by the Agency's Information Quality Guidelines but which it apparently has thus far declined to conduct.

A question arises concerning what to do if data for each maneuver cannot be obtained because, for example, the researchers discarded it or failed to record it.[90] For each study in which EPA can obtain only single values instead of each maneuver, EPA could perform a Monte Carlo simulation utilizing a range of alternative standard deviations such as those produced by Vaughan et al. (1989). We know that EPA has the raw data from Adams (2006), because Adams provided it to EPA on request and EPA utilized it (or at least portions of it) to produce its controversial reanalysis purporting to show a statistically significant FEV_1 decrement at 60 ppb (Brown 2007a, 2007b). It would be a simple matter for EPA to conduct the same analysis again in a Monte Carlo format with inter-maneuver variability incorporated. Obviously important questions include:

- Is the FEV_1 decrement that EPA staff previously observed at 60 ppb still statistically significant?
- How small must the inter-maneuver standard deviation be to achieve statistical significance?
- Is this standard deviation plausible in the real world?

The public surely would benefit from this analysis. Administrator Jackson also deserves to know the answers to these questions before making a final decision that, to date, hinges so crucially on an analysis that is demonstrably misleading even if every other complaint about it is ignored.

Table III:
Standard Deviations of Maneuvers by Range of Percent Predicted Value

	Range of Percent Predicted				
Maneuver	50	50-75	75-100	100	*Overall*
FEV_1 (JPF)	3.30	3.02	3.08	2.08	3.01
PEFR (JPF)	5.47	7.33	8.10	6.40	7.20
PEFR (WPF)	4.67	6.08	5.14	4.35	5.12

Source: Vaughan et al. (1989, p. 560, Table 2).
JPF: Jones Pulmonar Spirometer; WPF: mini-Wright peak flow meter.

[90] This would be a violation of the ATS technical guidelines, which require retention of "at a minimum" three satisfactory maneuvers. *See* Miller, Hankinson et al. (2005, p. 325).

IV. NONRESPONSE BIAS IN OBSERVATIONAL EPIDEMIOLOGY STUDIES

Several of the observational epidemiology studies on which Administrator Johnson relied in 2008, and which Administrator Jackson is relying today, involve samples that have potentially serious nonresponse bias. Nonresponse bias arises when a representative sample is selected but some choose to drop out of the study or not to participate at all. Nonparticipants and dropouts should not be assumed to have the same characteristics as those who participate or stay. Oftentimes, nonrespondents are a very different subsample.[91]

OMB has for decades maintained government-wide statistical policy standards and guidelines related to the management of nonresponse bias.[92] The standards are mandatory; the guidelines are best practices for achieving them. They are excerpted in Table IV.

These standards apply whenever an agency conducts or sponsors a collection of information, or through the operation of the IQA, whenever an agency disseminates information that a reasonable person would construe as reflecting the agency's endorsement. All of the literature on which EPA relies is thus covered indirectly, and several of these studies were conducted or sponsored by EPA or another Federal agency. This includes Gent et al. (2003, [NIEHS]), Korrick et al. (1998, [NIEHS, EPA, NIH]), and Mortimer et al. (2002, [EPA]).

Nonresponse bias analysis is required in any case where an item response rate falls below 70% or a unit response rate falls below 80%.[93] Gent et al. (2003, p. 1860) used a cohort of 1,002 infants, 357 they deemed to be eligible and 272 participated (76%). Given the size of the cohort, the true response rate is unclear. Mortimer et al. (2002, p. 701) used a sample of 846 children from a cohort of 1,528 enrolled in a multicity asthma study.[94] The response rate thus was 55% (though they describe it as "[a]pproximately 60%" [p. 700]). Both research teams performed statistical analyses in which it is assumed that the samples

---

[91] Not all samples are representative. For example, convenience samples are popular in epidemiology but they have no known sample properties and their results cannot be generalized to any known population. A relevant example in the ozone literature is Korrick et al. (1998), on which EPA relied on for the 2008 Rule and proposes to rely on again for the Reconsideration.

[92] The most recent edition is Office of Management and Budget (2006).

[93] *See* Office of Management and Budget (2006):

> "**Nonresponse bias** occurs when the observed value deviates from the population parameter due to differences between respondents and nonrespondents. Nonresponse bias may occur as a result of not obtaining 100 percent response from the selected cases" (p. 33).
>
> "**Item nonresponse** occurs when a respondent fails to respond to one or more relevant item(s) on a survey" (p. 31).
>
> "**Unit nonresponse** occurs when a respondent fails to respond to all required response items (i.e., fails to fill out or return a data collection instrument)" (p. 35).

[94] Mitchell et al. (1997, p. 238).

were representative (an assumption neither research team defended) and that their response rates were 100% (which is demonstrably wrong). Korrick et al. (1998) used a convenience sample that has no known sample properties. A total of 766 hikers volunteered to participate, 595 (78%) of whom provided spirometry data both before and after the hike. Statistical analyses were performed assuming that the convenience sample was representative and the response rate was 100%.

We are unaware of any nonresponse bias analyses published by these research teams, and the 2006 Criteria Document does not report that any were performed. It thus appears that none of these studies met minimum Federal statistical standards. A review of the 2006 Criteria Document indicates that the Agency has not conducted a nonresponse bias analysis for any study on which it relies. Thus, EPA noncompliance with Federal statistical policy appears to be much more substantial than we have documented here.

Table IV: Federal Statistical Policy Guidance Related to Nonresponse Bias (Excerpts)

Section 3.2 Nonresponse Analysis and Response Rate Calculation
Standard 3.2: Agencies must appropriately measure, adjust for, report, and analyze unit and item nonresponse to assess their effects on data quality and to inform users. Response rates must be computed using standard formulas to measure the proportion of the eligible sample that is represented by the responding units in each study, as an indicator of potential nonresponse bias.
The following guidelines represent best practices that may be useful in fulfilling the goals of the standard:
Guideline 3.2.1: Calculate all response rates unweighted and weighted. Calculate weighted response rates based on the probability of selection or, in the case of establishment surveys, on the proportion of key characteristics that is represented by the responding units. Agencies may report other response rates in addition to those given below (e.g., to show the range of response rates given different assumptions about eligibility) as long as the rates below are reported and any additional rates are clearly defined.
Guideline 3.2.2: Calculate unweighted unit response rates (RRU) as the ratio of the number of completed cases (or sufficient partials) (C) to the number of in-scope sample cases.
Guideline 3.2.3: Calculate weighted unit response rates (RRW) to take into account the different probabilities of selection of sample units, or for economic surveys, the different proportions of key characteristics that are represented by the responding units.
Guideline 3.2.4: Calculate the overall unit response rates for cross-sectional sample surveys (RROC) as the product of two or more unit-level response rates when a survey has multiple stages.
Guideline 3.2.5: Calculate longitudinal response rates for each wave. Use special procedures for longitudinal surveys where previous nonrespondents are eligible for inclusion in subsequent waves.

Section 3.2 Nonresponse Analysis and Response Rate Calculation
Guideline 3.2.6: Calculate item response rates (RRI) as the ratio of the number of respondents for whom an in-scope response was obtained (Ix for item x) to the number of respondents who were asked to answer that item.
Guideline 3.2.7: Calculate the total item response rates (RRT^x) for specific items as the product of the overall unit response rate (RRO) and the item response rate for item x (RRI^x)
Guideline 3.2.8: When calculating a response rate with supplemented samples, base the reported response rates on the original and the added sample cases. However, when calculating response rates where the sample was supplemented during the initial sample selection (e.g., using matched pairs), calculate unit response rates without the substituted cases included (i.e., only the original cases are used).
Guideline 3.2.9: Given a survey with an overall unit response rate of less than 80 percent, conduct an analysis of nonresponse bias using unit response rates as defined above, with an assessment of whether the data are missing completely at random.
Guideline 3.2.10: If the item response rate is less than 70 percent, conduct an item nonresponse analysis to determine if the data are missing at random at the item level for at least the items in question, in a manner similar to that discussed in Guideline 3.2.9.
Guideline 3.2.11: In those cases where the analysis indicates that the data are not missing at random, the amount of potential bias should inform the decision to publish individual items.
Guideline 3.2.12: For data collections involving sampling, adjust weights for unit nonresponse, unless unit imputation is done. The unit nonresponse adjustment should be internally consistent, based on theoretical and empirical considerations, appropriate for the analysis, and make use of the most relevant data available.
Guideline 3.2.13: Base decisions regarding whether or not to adjust or impute data for item nonresponse on how the data will be used, the assessment of nonresponse bias that is likely to be encountered in the review of collections, prior experience with this collection, and the nonresponse analysis discussed in this section. When used, imputation and adjustment procedures should be internally consistent, based on theoretical and empirical considerations, appropriate for the analysis, and make use of the most relevant data available. If multivariate analysis is anticipated, care should be taken to use imputations that minimize the attenuation of underlying relationships.
Guideline 3.2.14: In the case of imputing longitudinal data sets, use cross-wave imputations or cross-sectional imputations.
Guideline 3.2.15: Clearly identify all imputed values on a data file (e.g., code them).
Source: Office of Management and Budget (2006).

V. EPA HAS DIMISSED AS INVALID IN OTHER NAAQS' CONTEXTS THE SAME DATA IT RELIED UPON IN THE OZONE NAAQS

NAM has previously noted (both above and in the 2008 RFR) that EPA uses science only to support predetermined policy decisions. The strongest evidence occurs when EPA relies on studies if they purport to show risk for one air pollutant but discards them when they show no risk for another. When confronted with this evidence through information quality petitions, EPA simply refuses to adhere to the law or its own guidelines and divines a post hoc rationalization.

A. EPA's interpretation of science depends on the regulatory outcome it seeks.

For ozone, EPA has stated repeatedly that it considers valid and reliable those studies that rely on self-administered and self-reported lung function tests. But EPA has stated with equal clarity that these same studies are unreliable for use in analogous contexts, most notably the nitrogen dioxide (NO_2) and sulfur dioxide (SO_2) NAAQS. There is no conceivable *scientific* justification for these opposing views. EPA's selective and conflicting use of these studies suggests that EPA likes these studies in the case of ozone because researchers claim to have discovered statistically significant effects for ozone, but the Agency dislikes them in the case of NO_2 and SO_2 because researchers found no such effects.

We raised this matter in the 2008 RFR:

> On August 30, 2007, about six weeks after finalizing [the scientific record] and publishing the proposed rule, the Agency separately distributed for public comment and CASAC review its draft Integrated Science Assessment for nitrogen oxides. Unsurprisingly, some of the same studies that are relevant to estimating human health risks from ozone also are relevant to estimating analogous risks from NO_x. Very surprisingly, however, in the NO_x ISA EPA says that pulmonary function test data are "notoriously" unreliable.

These studies were so unreliable that EPA declined to use them. In the Draft NO_x ISA,

> EPA summarizes—and dismisses—several studies in which pulmonary function data were collected. Among them: the study by Mortimer et al. (2002), the same study of asthmatic children that, in the ozone Staff Paper, EPA said "suggest[s] that O_3 exposure may be associated with clinically significant changes in PEF in asthmatic children" and identified "plausible biological mechanisms that would

> explain delayed effects consistent with the distributed lag models that yielded that only statistically significant results."
>
> In the ozone Staff Paper, EPA considers the use of PEFR monitors by Mortimer et al. (2002) to be state of the art and their results persuasive.[95]

These opposite interpretations of the *same studies* cannot be justified by an appeal to science. They undoubtedly reflect different opinions about whether these studies support the *policy views* of EPA staff.

Because we pointed out the fundamental inconsistency in EPA's reasoning, we expected that the Agency would make changes in the Final NO_x ISA to eliminate it. That is exactly what EPA has done; the text in the Draft ISA that provided a transparent account of the inconsistency was deleted in the Final.[96]

In the 2008 RFR we pointed out that this was not an isolated instance in which EPA has interpreted science contingent on whether it supports staff policy views:

> In its Response to Comments, EPA is dismissive of the randomized panel study of asthmatic children by Schildcrout et al. (2006). EPA faulted it for having just 990 subjects. "As a result," EPA writes, "the total number of children observed by Schildcrout et al. is not comparable to other large multi-city studies that examined the effect of O_3 concentrations on asthma exacerbation, such as Mortimer et al. (2002)." This is an especially odd complaint, inasmuch as the study by Mortimer et al. (2002) included 846 children.
>
> EPA's low opinion of Schildcrout et al. (2006) is limited to ozone, however. In EPA's final Integrated Science Assessment for SO_2, EPA says "the strongest epidemiological evidence for an association between respiratory symptoms and exposure to ambient and SO_2 comes from two large multi-city studies"—Mortimer et al. (2002) and Schildcrout et al. (2006). The difference is that Schildcrout et al. (2006) reported a statistically significant positive association between SO_2 and respiratory symptoms, but no association with ozone. EPA likes Mortimer et al. (2002) for both ozone and SO_2; Mortimer et al. (2002) found positive associations for both.[97]

---

[95] National Association of Manufacturers (2008, pp. 58-59, internal references omitted).
[96] U.S. Environmental Protection Agency (2008a).
[97] National Association of Manufacturers (2008, pp. 59-60, internal footnote and references omitted).

B. EPA is unwilling to address inconsistencies in its interpretation of science.

EPA's refusal to respond to this aspect of the 2008 RFR is consistent with its refusal to respond in other contexts. NAM identified this same information quality defect in a Request for Correction filed in June 2009 on the final NO_x Integrated Science Assessment ("2009 NO_x RFC").[98] In its reply, EPA simply refused to respond to the substance of the issue. Contrary to any procedure set forth in the EPA IQG, the Agency treated the RFC as a public comment on a different information dissemination—in this case, a proposed rule that was published *after* the RFC was submitted.[99] NAM then filed a Request for Reconsideration ("2009 NO_x RFR"),[100] to which EPA replied on February 16, 2010—seven days after promulgating a final rule based substantially on the document containing the information quality errors identified in the 2009 NO_x RFC.[101] Of course, the point of filing an RFR was to help avert a situation in which fatal information quality defects would provide the foundation for a major regulatory decision.

Unfortunately, EPA's apparent strategy in responding to these requests is delay. In responding to the 2009 RFR, EPA states that it is delaying a response "[d]ue to the complexity of the issues raised."[102] But the RFC did not raise a complex issue; indeed, it is hard to imagine a simpler information quality defect than the adoption of opposite interpretations of the same scientific study in two different regulatory contexts. Nonetheless, EPA has decided to postpone a genuine response to the 2009 NO_x RFR at least until May 16, 2010. By that time, the revised NO_2 NAAQS promulgated on February 9 may be challenged, thereby providing EPA with a possible excuse for a further delay in responding.

VI. EPA CONTINUES TO FAIL TO DISTINGUISH BETWEEN CASAC'S SCIENTIFIC REVIEW AND ITS POLICY RECOMMENDATIONS

As we described in the 2007 RFC and 2008 RFR, CASAC has an admittedly complicated role under the Clean Air Act. It is supposed to provide the Administrator with a review of the scientific database that is both objective and independent of the EPA staff, but also to offer policy recommendations to the Administrator. As we noted in great detail in the 2007 RFC and 2008 RFR, CASAC clearly had trouble keeping these two functions distinct. EPA's actions were singularly unhelpful in this regard, and that has resulted in myriad information quality defects in the way the Agency has utilized CASAC's input.

---

[98] National Association of Manufacturers (2009a).
[99] Kadeli (2009).
[100] National Association of Manufacturers (2009b)
[101] U.S. Environmental Protection Agency (2010b).
[102] Cheatham (2010).

These defects are exacerbated in the Proposed Reconsideration. The 2007 RFC and 2008 RFR explained why CASAC's science review and policy advice could not be disentangled. EPA did not ask CASAC to keep them separate—indeed, EPA never said anything to CASAC about the IQA or EPA's commitment to apply information quality principles throughout its operations. The 2007 RFC and 2008 RFR also explained why the IQA and applicable guidelines required EPA to make a good faith effort to disentangle science and policy in CASAC's various letters, and noted that EPA had failed to do so.

Nonetheless, Administrator Johnson at least appears to have been well aware of the problem, as the Final Rule makes clear in its description of CASAC's input:

> With respect to CASAC's recommended range of standard levels, the Administrator observes that the basis for its recommendation appears to be a mixture of scientific and policy considerations.[103]

Caveats such as this are missing from the Proposed Reconsideration. A reasonable inference is that EPA does not want to admit that CASAC's *scientific* review lacks objectivity because it is suffused with the *policy judgments* of its members. This poses a problem because Administrator Jackson wants to claim that *science* requires a primary standard lower than 75 ppb, and CASAC is needed to provide that scientific support.

A. Clear distinctions between science and policy judgment are a hallmark of Federal risk management policy.

In 1983, a committee of the National Research Council established a fundamental principle that has grounded U.S. risk management policy ever since:

> We recommend that regulatory agencies take steps to establish and maintain a clear conceptual distinction between assessment of risks and consideration of risk management alternatives; that is, the scientific findings and policy judgments embodied in risk assessments should be explicitly distinguished from the political, economic, and technical considerations that influence the design and choice of regulatory strategies.[104]

This *distinction* was never intended to imply a *separation* of risk assessment from risk management, although that is how EPA first implemented it.[105] It was intended to ensure clarity, so that scientific matters were left to scientists and public officials made policy decisions, with neither group interfering in the other's rightful responsibilities.

---

[103] U.S. Environmental Protection Agency (2008b, p. 16482).
[104] National Research Council (1983, p. 7).
[105] North (2003).

Over the years, EPA has repeatedly expressed its institutional support for this principle. That rhetorical support has not always been ratified by practice, however. For example, the EPA staff has vigorously defended its bureaucratic prerogative to make crucial policy decisions under the cover of ostensibly scientific risk assessment.[106] The 2007 RFC and 2008 RFR document a long list of instances in which the EPA staff used the cover of science to arrogate decision-making authority reserved by law to the Administrator. EPA administrators' practical ability to exercise lawful policy judgment is profoundly affected by the extent to which their policy views are aligned with those of the Agency staff.

B. CASAC has been diminished by EPA's failure to provide it with effective guidance about maintaining a clear distinction between science and policy.

EPA made CASAC's job immeasurably more difficult by failing to inform the panel about applicable information quality principles and practices, and by failing to even ask the panel to maintain a clear distinction between its scientific review and its policy advice. The written materials and transcripts of in-person meetings show that EPA staff from the Office of Air and Radiation, the Office of Research and Development, and the Science Advisory Board never alerted CASAC to the Agency's information quality guidelines. The SAB staff responsible for coordinating the CASAC review seems to have just ignored what the Agency's Peer Review Handbook says on the subject.[107]

We also pointed out in the 2007 RFC and the 2008 RFR that EPA's entire regulatory development process for the ozone NAAQS revision was bereft of any attention to information quality. There are not even throwaway boilerplate references in the Criteria Document, the Staff Paper, or any other document containing influential information subject to information quality principles, practices, and standards. This contravened the Agency's express written commitment, made in 2002, to incorporate information quality principles and practices throughout its operations.[108] The first time EPA ever dealt with information quality occurred when we submitted the 2007 RFC. As the 2008 RFR makes clear, EPA's response was incomplete, troubling in its evasiveness, and misleading. We

---

[106] U.S. Environmental Protection Agency Office of the Science Advisor (2004).

[107] *See* National Association of Manufacturers (2008, p. 12) and U.S. Environmental Protection Agency (2006b, pp. 16-18). In the 2008 RFR, we called a specific version of this phenomenon as the Iron Law of EPA Staff Ozone Health Risk Assessment and Characterization (2008, pp. 13-14).

[108] In its information quality guidelines, EPA implied that this would be simple because the Agency had achieved the Information Quality Act's purposes before it was enacted. Notice the use of present tense: "EPA ensures and maximizes the quality of the information we disseminate by implementing well established policies and procedures within the Agency as appropriate to the information product. There are many tools that the Agency uses such as the Quality System, review by senior management, peer review process, communications product review process, the web guide, and the error correction process" (U.S. Environmental Protection Agency 2002, p. 19, internal footnotes omitted).

specifically asked EPA to provide clarity concerning which inputs from CASAC it was interpreting as science and which as policy advice. Because EPA has not replied to the 2008 RFR, the Agency has not fulfilled its administrative duties with respect to any of these information quality errors.[109]

C. CASAC's recommendations are undermined by its failure to distinguish appropriately between science and policy.

With this history it is not surprising that CASAC was confused by its two distinctive roles and that this confusion was exacerbated when Administrator Johnson made his 2008 decision. On its own accord, CASAC produced and sent to Administrator Johnson an unsolicited letter strenuously objecting to his decisions.[110] CASAC's confusion is obvious in certain parts of this letter:

> It is the Committee's consensus scientific opinion that your decision to set the primary ozone standard above this range fails to satisfy the explicit stipulations of the Clean Air Act that you ensure an adequate margin of safety for all individuals, including sensitive populations.[111]

Not only was CASAC unable to see the distinction between *objective* scientific review and *subjective* policy recommendations, it claimed to have legal expertise that gave it a superior ability to interpret the law.

Of course, the "explicit stipulations" to which CASAC refers are policy judgments, not science, the presumptive domain of CASAC members' expertise. There is no scientific definition for "margin of safety"; indeed, even the term "safety" cannot be defined scientifically. Nor are there scientific definitions for what margin of safety is "adequate" or what constitutes a "sensitive subpopulation." These are legal terms of art in the Clean Air Act; outside the Clean Air Act they have no meaning. The Administrator is legally required to allow science to inform his policy judgment, but if the law intended for science to dictate decision-making, these nonscientific factors would have been absent.

While it is true that EPA provided no assistance in distinguishing between science and policy, CASAC undermined its own scientific credibility by failing to provide this distinction. No reader of CASAC's reports—including the EPA Administrator— can clearly

---

[109] As the 2007 RFC and the 2008 RFR, make clear CASAC is not an "agency" as defined in 44 U.S.C. § 3502(1). Thus, it is exempt from the Information Quality Act and its implementing guidelines. EPA, of course, is not exempt, and how it manages information provided by CASAC is clearly covered. EPA cannot disseminate representations of fact or knowledge it obtains from CASAC and merely presume that it meets applicable information quality standards for utility, integrity and objectivity.

[110] Henderson (2008).

[111] Henderson (2008, p. 2).

distinguish its scientific content from its policy advice, or be sure that what appears to be scientific content is expressed objectively, as the Clean Air Act sets forth as CASAC's primary mission.

The Proposed Reconsideration multiplies these problems. In the preamble to the Proposed Reconsideration, EPA cites CASAC's policy advice in a way that is clearly intended to convey the impression that it is actually science.[112] In essence, EPA is attempting to rely on CASAC credibility to support a different policy choice. This is the same lack of presentational objectivity that we noted in the 2008 RFR.

In the Proposed Reconsideration, however, there is a new and more egregious information quality error. EPA seeks to rebrand CASAC's policy advice as science to evade public accountability for making a decision that is based almost entirely on policy considerations. Instead of transparently stating that Administrator Jackson disagrees with Administrator Johnson's policy decision, EPA is recharacterizing CASAC's opinions as "science" so that EPA can imply the science is compelling EPA's reconsideration of the 2008 standard. In addition to remedies sought in the RFR, this RFC seeks the following specific remedy with respect to the way the Proposed Reconsideration treats inputs from CASAC. In every instance where EPA cites a CASAC statement as "science," it should document that:

- The statement is not a policy judgment; it contains only representations of facts or knowledge and thus is capable of being refuted upon the application of data and analysis.
- The statement is substantively objective; it has no perceptible inaccuracies or biases, such as an embedded or unstated preferences concerning what standard ought to be set.
- The statement is presentationally objective; it is presented in an accurate, clear, complete, and unbiased manner.

Like the preamble to the 2008 Final Rule, the preamble to the Reconsideration makes no such showing.

To date, EPA has simply failed to comply with information quality principles or to adhere to its own policies and commitments. We are hopeful that the Agency will now engage in a real "reconsideration," and follow the requirements of the law.

---

[112] *See* U.S. Environmental Protection Agency (2010a, p. 2992, citing the same portion of CASAC's April 7, 2008 letter).

VII. REFERENCES

Adams WC. 2006. Comparison of Chamber 6.6-hr Exposures to 0.04-0.08 ppm Ozone via Square-Wave and Triangular Profiles on Pulmonary Responses. *Inhalation Toxicology* 18: 127-136.

Brown JS. 2007a. *The effects of ozone on lung function at 0.06 ppm in healthy adults (EPA-HQ-OAR-2005-0172-0175)*. Research Triangle Park, N.C.: U.S. Environmental Protection Agency National Center for Environmental Assessment.

Brown JS. 2007b. *The effects of ozone on lung function at 0.06 ppm in healthy adults (January 2007 draft)*. U.S. Environmental Protection Agency National Center for Environmental Assessment.

Brown JS, Bateson TF, McDonnell WF. 2008. Effects of Exposure to 0.06 ppm Ozone on FEV_1 in Humans: A Secondary Analysis of Existing Data. *Envionmental Health Perspectives* 116(8): 1023-1026.

Cheatham R. 2010. *Response to Request for Reconsideration (#090002A) on EPA's Response to National Association of Manufacturers Request for Correction: Integrated Science Assessment--Oxides of Nitrogen (#09002)*. Available: http://www.epa.gov/quality/informationguidelines/documents/09002A-interim.pdf [accessed March 20, 2010].

Gent MH, Triche EW, Holford TR, Belanger K, Bracken WS, Beckett MR, Leaderer BP. 2003. Association of Low-level Ozone and Fine Particles with Respiratory Symptoms in Children with Asthma. *Journal of the American Medical Association (JAMA)* 290: 1859-1867.

Henderson R. 2008. *April 7, 2008 Letter to Stephen L. Johnson from CASAC on 'Clean Air Scientific Advisory Committee Recommendations Concerning the Final Rule for the National Ambient Air Quality Standards for Ozone'*. "CASAC April 7, 2008 Letter on O_3 NAAQS". Washington, D.C.: U.S. Environmental Protection Agency Office of the Science Advisory Board.

Jackson LP. 2009. *Memo to EPA Employees: Transparency in EPA's Operations*. Available: [accessed March 13, 2010].

Kadeli L. 2009. *Response to National Association of Manufacturers Request for Correction: Integrated Science Assessment--Oxides of Nitrogen*. Available: http://www.epa.gov/quality/informationguidelines/documents/09002-response.pdf [accessed March 18, 2010].

Korrick SA, Neas LM, Dockery DW, Gold DR, Allen GA, Hill LB, Kimball KD, Rosner BA, Speizer FE. 1998. Effects of Ozone and Other Pollutants on the Pulmonary Function of Adult Hikers. *Environmental Health Perspectives* 106: 93-99.

McDonnell WF, Stewart PW, Smith MV. 2007. The Temporal Dynamics of Ozone-induced FEV_1 Changes in Humans: An Exposure-response Model. *Inhalation Toxicology* 19(6-7): 483-494.

Miller MR, Crapo R, Hankinson J, Brusasco V, Burgos F, Casaburi R, Coates A, Enright P, van der Grinten CPM, Gustafsson P, Jensen R, Johnson DC, MacIntyre N, McKay R, Navajas D, Pedersen OF, Pellegrino R, Viegi G, Wanger J. 2005. General considerations for lung function testing. *European Respiratory Journal* 26(1): 153-161.

Miller MR, Hankinson J, Brusasco V, Burgos F, Casaburi R, Coates A, Crapo R, Enright P., van der Grinten CPM, Gustafsson P, Jensen R, Johnson DC, MacIntyre N, McKay R, Navajas D, Pedersen OF, Pellegrino R, Viegi G, Wanger j. 2005. Standardisation of spirometry. *European Respiratory Journal* 26(2): 319-338.

Mitchell H, Senturia Y, Gergen P, Baker D, Joseph C, McNiff-Mortimer K, Wedner HJ, Crain E, Eggleston P, Evans Rr, Kattan M, Kercsmar C, Leickly F, Malveaux F, Smartt E, Weiss K. 1997. Design and Methods of the National Cooperative Inner-City Asthma Study. *Pediatric Pulmonology* 24(4): 237-252.

Mortimer KM, Neas, L.M.,, Dockery DW, Redline S, Tager IB. 2002. The Effect of Air Pollution on Inner-city Children with Asthma. *The European Respiratory Journal* 19(4): 699-705.

National Association of Manufacturers. 2007. *Request for Correction under EPA's Information Quality Guidelines; Letter and Comments Sent to Molly A. O'Neill (Assistant Administrator, Office of Environmental Information and Chief Information Officer) re: Proposed Rule – National Ambient Air Quality Standards for Ozone; Docket No. OAR-2005-0172-4275*. Available: [accessed November 5, 2008].

National Association of Manufacturers. 2008. *Request for Reconsideration under EPA's Information Quality Guidelines: Letter and Comments Sent to Molly A. O'Neill (Assistant Administrator, Office of Environmental Information and Chief Information Officer) re: Proposed Rule – National Ambient Air Quality Standards for Ozone; Docket No. OAR-2005-0172-4275 [October 14, 2008]*. Available: [accessed November 5, 2008].

National Association of Manufacturers. 2009a. *Request for Correction: Integrated Science Assessment--Oxides of Nitrogen*. Available: http://www.epa.gov/quality/informationguidelines/documents/09002.pdf [accessed March 18, 2010].

National Association of Manufacturers. 2009b. *Request for Reconsideration of RFC #09002--Oxides of Nitrogen*. Available: http://www.epa.gov/quality/informationguidelines/documents/09002A.pdf [accessed March 18, 2010].

National Research Council. 1983. Risk Assessment in the Federal Government: Managing the Process. Washington, D.C.: National Academies Press.

North DW. 2003. Reflections on the Red/Mis-Read Book, 20 Years After. *Journal of Human and Ecological Risk Assessment* 9(5).

Obama B. 2009a. Memorandum of January 21, 2009: Freedom of Information Act. *Federal Register* 74(15): 4683-4684.

Obama B. 2009b. Memorandum of January 21, 2009: Transparency and Open Government. *Federal Register* 74(15): 4685-4686.

Office of Management and Budget. 1999. *Circular A-110: Uniform Administrative Requirements for Grants and Agreements With Institutions of Higher Education, Hospitals, and Other Non-Profit Organizations*. "OMB Circular A-110". Washington, D.C.

Office of Management and Budget. 2002. Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated by Federal Agencies; Notice; Republication. *Federal Register* 67(36): 8452-8460.

Office of Management and Budget. 2006. *Standards and Guidelines for Statistical Surveys*. Available: [accessed October 1, 2007].

Pellegrino R, Viegi G, Brusasco V, Crapo RO, Burgos F, Casaburi R, Coates A, van der Grinten CPM, Gustafsson P, Hankinson J, Jensen R, Johnson DC, MacIntyre N, McKay R, Miller MR, Navajas D, Pedersen OF, Wanger J. 2005. Interpretative strategies for lung function tests. *European Respiratory Journal* 26: 948-968.

Schelegle ES, Morales CA, Walby WF, Marion S, Allen RP. 2009. 6.6-hour inhalation of ozone concentrations from 60 to 87 parts per billion in healthy human. *American Journal of Respiratory and Critical Care Medicine* 180(3): 265-272.

Schildcrout JS, Sheppard L, Lumley T, Slaughter JC, Koenig JQ, Shapiro GG. 2006. Ambient air pollution and asthma exacerbations in children: an eight-city analysis. *American Journal of Epidemiology* 164(6): 505-517.

U.S. Environmental Protection Agency. 2002. *Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated by the Environmental Protection Agency (EPA/260R-02-008)*. "EPA IQG".

U.S. Environmental Protection Agency. 2006a. *Air Quality Criteria for Ozone and Related Photochemical Oxidants: Volume I of III (EPA 600/R-05/004aF)*. "2006 O_3 NAAQS CD Vol I". Research Triangle Park, N.C.: U.S. Environmental Protection Agency Office of Research and Development National Center for Environmental Assessment-RTP Office.

U.S. Environmental Protection Agency. 2006b. *Peer Review Handbook*. "EPA Peer Review Handbook 3rd Ed". Washington, D.C.: U.S. Environmental Protection Agency Science Policy Council.

U.S. Environmental Protection Agency. 2007. *Draft Integrated Science Assessment for Oxides of Nitrogen — Health Criteria; August 2007*. "2007 1st Draft NO_x ISA". Washington, D.C.: U.S. Environmental Protection Agency Office of Air Quality Planning and Standards.

U.S. Environmental Protection Agency. 2008a. *Final Integrated Science Assessment for Oxides of Nitrogen — Health Criteria (EPA/600/R-08/071)*. "2008 Final NO_x ISA". Washington, D.C.: U.S. Environmental Protection Agency Office of Air Quality Planning and Standards.

U.S. Environmental Protection Agency. 2008b. National Ambient Air Quality Standards for Ozone; Final Rule. *Federal Register* 73(60): 16436-16514.

U.S. Environmental Protection Agency. 2008c. *Responses to Significant Comments on the 2007 Proposed Rule on the National Ambient Air Quality Standards for Ozone (July 11, 2007; 72 FR 37818); Docket ID EPA-HQ-OAR-2005-0172-7152.1*. "2008 O_3 NAAQS Response to Comments" (U.S. Environmental Protection Agency, ed). Washington, D.C.: U.S. Environmental Protection Agency

U.S. Environmental Protection Agency. 2009a. *Letter from Robert J. Meyers, Principal Deputy Assistant Administrator for Air and Radiation, U.S. Environmental Protection Agency, to Bryan J. Brendle, Director, Energy and Resource Policy, National Association of Manufacturers*. Washington, D.C.

U.S. Environmental Protection Agency. 2009b. *Provisional Assessment of Recent Studies on Health and Ecological Effects of Ozone Exposure [EPA/600/R–09/101, Docket ID EPA-HQ-OAR-2005-0172-7232]*. Research Triangle Park: National Center for Environmental Assessment, 55 pp.

U.S. Environmental Protection Agency. 2010a. National Ambient Air Quality Standards for Ozone; Proposed Rule. *Federal Register* 75(11): 2938-3052.

U.S. Environmental Protection Agency. 2010b. Primary National Ambient Air Quality Standards for Nitrogen Dioxide; Final Rule. *Federal Register* 75(26): 6474-6537.

U.S. Environmental Protection Agency Office of the Science Advisor. 2004. *An Examination of EPA Risk Assessment Principles and Practices; Staff Paper, EPA/100/B-04/001*. Available: [accessed August 30, 2005].

Vaughan TR, Weber RW, Tipton WR, Nelson HS. 1989. Comparison of PEFR and FEV1 in Patients with Varying Degrees of Airway Obstruction; Effect of Modest Altitude. *Chest* 95: 558-562.