

Richard B. Belzer, Ph.D.  
President  
[Belzer@RegulatoryCheckbook.org](mailto:Belzer@RegulatoryCheckbook.org)



September 19, 2002

Ms. Kim Nelson  
Chief Information Officer  
U.S. Environmental Protection Agency  
Washington, DC 20450

DOCKET ID OEI-10014

Dear: Ms. Nelson:

I am pleased to provide limited comments on the September 6<sup>th</sup> external review draft of EPA's latest information quality guidance document entitled *Assessment Factors for Evaluating the Quality of Information from External Sources*. This document was publicly noticed in the Federal Register only on September 9<sup>th</sup> (67 FR 57225-57226). Somewhat remarkably, EPA initially provided just 11 days for public comment. I understand that EPA has extended this period another 10 days.

Some public commenters may be pleased to have the additional time. I do not share their enthusiasm. Both the initial 11-day public comment period and its 21-day successor are clearly inadequate and unreasonable. EPA cannot credibly argue that it is both constrained by the impending October 1<sup>st</sup> for issuance of final agency guidelines and simultaneously assert that it intends to pay any attention to the comments it receives by midnight on September 30<sup>th</sup>. By the time the public comment period expires, EPA's draft final information quality guidelines will be on display at the Office of the Federal Register. This will surely make it difficult for EPA to make timely revisions.

Therefore, the comments herein are provided largely for the administrative record and to benefit other members of the public interested in EPA's implementation of the Data Quality Act. Yet they are provided with some trepidation because especially pernicious outcomes can result when public commenters are purposefully kept uninformed. There is every reason to be deeply concerned that EPA might make changes nominally based on these comments that are fundamentally at odds with my intent.

## **PROCEDURAL CONCERNS**

The external review draft raises a number of procedural concerns besides the matter of timing.

Why is EPA seeking public comment on a guidance document that modifies final guidelines that it has not yet issued? Informed public comment is not possible because EPA's final guidelines will not be published until October 1<sup>st</sup>. Commenters seem to be left to interpret the document based on OMB's government-wide directive or EPA's proposed guidelines. The former is possible but yields little fruit, for hardly anything in this document actually implements any significant element of OMB's directive. Interpreting this document based on EPA's proposed guidelines is possible but frustrating because that document did not reflect the high level of commitment that you and other senior EPA officials had promised. In the cover letter to my public comments on the proposed guidelines, I wrote:

EPA's proposal abuses the opportunity for implementation flexibility by instead striving to evade all reasonable information quality standards. In many places, EPA's efforts are so brazen as to constitute self-parody. It is extremely difficult to interpret these departures as anything other than systematic evidence of bad faith.

With respect to OMB's directive that EPA craft standards for third-party information, I stated:

EPA proposes to indefinitely postpone the development of information quality guidelines with respect to third-party data. EPA's proposal merely punts this issue to an undetermined future date. The public has nothing on which to comment.

In principle, publication of this document finally gives the public something to comment upon. But prospective commenters lack the foundation document which this document modifies because EPA has not released it. EPA acts like a blackjack dealer who claims that he has "21" but refuses to display the card hidden under the ace.

How does this document integrate with EPA's final information quality guidelines? Despite the fact that this is a critical issue, it cannot be reliably ascertained because EPA provides no information relevant to the point. A reasonable inference is that EPA plans to impose higher standards on (at least some) third-party data than it intends to abide by in its own data collection.

#### **INFORMATION POLICY CONCERNS**

Does EPA intend to use high information quality standards as a shield against third-party information that, for whatever reason, it does not like? A reasonable inference is that EPA intends to establish higher standards for third-party data than it subjects upon its own data collections. Such a plan would be contrary to both law and OMB's government-wide implementing directive. Indeed, nothing in OMB's documents authorizes or encourages agencies to establish differential information quality standards that penalize third parties. The fundamental purpose of both the law and OMB's government-wide directive is to protect the public from governmental dissemination or use of substandard information. They were not intended to protect an agency from superior quality information from nongovernmental sources.

This inference could be completely wrong. It can be overcome only by persuasive evidence that its final information quality guidelines – as well as any document it references directly (such as this one) or indirectly (such as those in Appendix 1) -- contain language clearly stating that information quality standards do not differ in any material respect depending on their origin. If EPA never intended to establish differential standards, it would be a simple matter to add declarative sentences to the document such as the following:

In all cases, third-party information must equal or exceed in quality whatever information EPA otherwise would disseminate or use for an equivalent purpose. Conversely, in no case must third parties satisfy information quality standards that EPA itself does not meet for an equivalent purpose.

What is EPA's factual basis for asserting that the documents it cites in Appendix 1 ("Representative Reference Materials") actually foster achievement of the objectives of the Data Quality Act and OMB's government-wide implementation directive? It may well be true that some (or perhaps all) of these documents are concerned with quality as an information attribute. However, none was published after the Data Quality Act was passed or after OMB issued its government-wide implementation directive. Both the law and OMB's directive encouraged agencies to use existing management systems to the extent practicable and consistent with information quality objectives. EPA seems to have completely misread OMB's directive, which states:

OMB encourages agencies to incorporate the standards and procedures required by these guidelines into their existing information resources management and administrative practices rather than create new and potentially duplicative or contradictory processes.

That is, EPA is welcome to *modify* its existing standards and procedures. Instead, EPA proposes to *modify* OMB's *directive* to accommodate its existing practices.

On this point the document is especially worrisome. As EPA did in its proposed information quality guidelines, EPA asserts (a) that each item in its panoply of management systems and guidance documents really was presciently designed to achieve the goals and objectives of the Data Quality Act, and (b) that the Agency actually complies with all these procedures and requirements. Assertions are not substitutes for evidence, however. If EPA intends to rely on these documents it still needs to audit them thoroughly and publicly to ascertain the extent to which they need to be amended. The assessment factors guidance is incomplete without such an evaluation if it references these other documents. If EPA's final information quality guidelines also make these assertions without persuasive evidence, then they will be fatally flawed as well.

Such an evaluation might reveal substantial uncertainty about what these documents actually say. Some of them are clear as mud. EPA guidance documents are often worded vaguely to preserve the Agency's interpretative discretion -- just as this draft guidance document was written. When what a document says is subject to dispute it may be impossible to know whether compliance with the document has occurred.

An audit also might reveal that actual Agency compliance with these documents is burdened by a checklist mentality. That is, substantive compliance may be incidental to procedural compliance. A checklist mentality would not be surprising for it is starkly displayed in both EPA's proposed information quality guidelines and this document: EPA has many guidance documents that mention information quality; ergo, the Agency achieves it.

Why do EPA's proposed assessment factors make the perfect the enemy of the good? The document is silent about the baseline level of information quality that third-party information would amend or supplant. Clearly, very high standards ought to apply in cases where EPA's own information exceeds all applicable information quality standards, including the high standard for influential information. But it would be a foolish third party who invests substantial resources in developing new information in these instances. It is much more likely (and quite nearly certain) that third parties will not do this. Rather, they will develop new information only in cases where EPA's baseline information does not meet applicable information quality standards and EPA's dissemination or use of substandard information causes significant and persistent harm.

EPA appears to be demanding that third parties satisfy informational perfection in order to be considered seriously. If true, this would be a deeply cynical way for EPA to resist unwelcome improvements in information quality. The standard of review for third-party information should be based on the quality of EPA information that it would supplant. Third-party information that is superior to EPA's baseline information should be unambiguously preferred even if it does not fully meet all information quality standards. A standard that penalizes third parties for imperfections less severe than those afflicting EPA's baseline information is clearly inappropriate and contrary to both law and OMB's government-wide directive.

EPA should graciously accept and fully utilize all third-party information that is superior in quality to whatever information the Agency otherwise would use or disseminate -- even if this new information does not fully satisfy all applicable information quality standards. The document ought to be amended to include an explicit commitment such as the following:

EPA will always accept, preferentially disseminate and use information from third parties that exceeds in quality the information that it otherwise would have disseminated or used for any equivalent or higher purpose. In no instance will EPA reject third-party information on the ground that it fails to fully achieve all desirable or applicable information quality standards.

## TECHNICAL CONCERNS

Several specific aspects of EPA's proposed assessment factors deserve comment. At first glance, these factors appear completely reasonable. Further reflection suggests, however, that they do not actually add much content. A case can be made that, on balance, they obfuscate more than they clarify.

Why is EPA proposing to invent these new terms of art when it has thus far failed to provide workable interpretations of the terms already established by OMB? EPA's first responsibility under the Data Quality Act is to flesh out agency-specific details to implement OMB's government-wide directive. Instead, EPA proposes to invent *more* new terms. Worse, EPA's language cannot be mapped to OMB's in any logical way. A reasonable inference is that EPA has decided to rewrite OMB's directive rather than implement it.

What do EPA's proposed assessment factors mean? Each of EPA's five factors has some merit. In several cases, however, EPA's description of the factor or how the Agency intends to apply it is confusing, inconsistent or out of place.

“Soundness”

EPA defines “soundness” as “reasonable for and consistent with the intended application” and “scientifically/technically appropriate.” On the surface this appears to be an elaboration of the basic quality standard for “sound science” that OMB issued in its government-wide directive. EPA's primary task in this area, however, was to “adopt or adapt” the basic standards of quality set forth in the Safe Drinking Water Act. The relevant language from SDWA requires the use of:

- (i) the best available, peer-reviewed science and supporting studies conducted in accordance with sound and objective scientific practices; and
- (ii) data collected by accepted methods or best available methods (if the reliability of the method and the nature of the decision justifies use of the data).

EPA says nothing about the SDWA standards, borrows one of its key terms, and then surrounds that term with murky legalistic language that is largely devoid of content. Because EPA has excellent skills at crafting clear language when it wants to, one must infer that the Agency's objective here is to maximize its interpretative discretion through semantic turbidity.

The bullets purporting to elaborate on the meaning of “soundness” either fail to do so or are fundamentally inconsistent with law and OMB's government-wide directive.

Bullet #1

Bullet #1 says that third-party information will be evaluated with respect to whether “procedures, measures, methods, or models employed to develop the information reasonable and consistent with sound scientific theory *or standard approaches?*” (emphasis added). Nowhere in either the SDWA language, or in OMB's government-wide implementing directive, is there a reference to *standard approaches* – much less any hint that they are desirable. Indeed, both the Data Quality Act and the relevant text of SDWA exist *because standard approaches often fail* to meet acceptable minimum quality standards.

Two prominent examples of *standard approaches* that do not meet the basic quality standard set forth in SDWA should make this concern clear. First, in risk assessment EPA routinely relies on *standard approaches* that cannot be characterized as “objective scientific practices.” These methods purposefully and methodically embed policy preferences into the estimation of risk. Put bluntly, they are intentionally biased. A case can be made for making certain decisions based on risk-averse preferences or an abundance of caution. However, no case can be made that purposefully and methodically biased methods constitute “objective scientific practices.”

Second, stated-preference models are increasingly used to estimate individuals' willingness-to-pay for environmental amenities and risk reduction. Such methods may be warranted because of the difficulty of obtaining data from revealed preference methods, especially in cases where no market transactions can be observed. Yet, in the vast majority of cases, stated preference methods fail to yield results that are consistent with elementary economic theory. Stated-preference methods may be the best we have because of inherent problems with direct measurement. Also, these methods have increasingly become part of analysts' toolbox of "standard approaches." In this case standardization really means familiarity, and it is inappropriate to presume that familiarity makes them "objective scientific practices."

The basic quality standard in SDWA includes language to deal with situations in which objective scientific practices are not available. These are termed "accepted methods or best available methods." Therefore, a much better approach is for EPA to craft clear definitions that highlight important distinctions among these concepts. Language such as the following would add needed clarity and faithfully implement both the law and OMB's government-wide directive:

"Soundness" means practices that are:

- (a) based on the scientific method;
- (b) founded on documented and disclosed research protocols that comply with high standards of scientific procedure and conduct;
- (c) blind with respect to investigators' ex ante knowledge of critical information, unless such prior knowledge is essential or unavoidable;
- (d) neutral with respect to public policy preferences or judgments; and
- (e) consistent with OMB's government-wide information quality guidelines and EPA's information quality guidelines.

"Accepted methods" mean:

- (a) methods that comply with the Paperwork Reduction Act and 5 C.F.R. 1320 where these authorities apply, and do not include any sample of nine or fewer;
- (b) approved test methods or techniques in any case where methods or techniques have been validated, and do not include unapproved test methods; and
- (c) compliance with generally accepted methods in the applicable field where neither (a) nor (b) apply.

"Best available methods" mean:

- (a) any sample of nine or fewer where such sample satisfies statistical criteria and tests showing that it is statistically valid and reliable under the circumstances for which the data are used;
- (b) unapproved test methods or techniques where such methods or techniques satisfy statistical criteria and tests showing that they are valid and reliable under the circumstances for which they are used.

If EPA established definitions such as these, it would “adopt or adapt” the basic quality standards set forth in the Safe Drinking Water Act and better inform third parties concerning what it means for information to be “sound.”

#### Bullet #2

Bullet #2 says that EPA will evaluate third-party information based on whether “novel or alternative theories or approaches are used,” and if so on how “clearly are they explained and the differences [with “standard approaches”?] highlighted?” Separately, EPA proposes “clarity and completeness” as an assessment factor. It is therefore not clear why EPA has chosen to embed language related to “clarity and completeness” within an assessment factor that is presumably unrelated.

One possible explanation is that EPA intends to erect a barrier against “novel or alternative theories or approaches.” This barrier would exceed the reasonable expectation that third parties explain “novel or alternative” approaches with “clarity and completeness,” for otherwise the language in bullet #2 is both redundant and misplaced. By including it here EPA establishes an inherent bias in favor of its own “standard approaches” and against scientific methods that it finds unduly creative.

The practical effect may be to thwart scientific progress that otherwise could occur in directions EPA finds uncomfortable. Should third parties fund and develop new, high-quality scientific methods for risk assessment or economic analysis or innovative applications of existing methods, EPA proposes to give itself additional tools to reject them – unless and until they become “standard approaches” at EPA.

#### Bullet #3

Bullet #3 specifies that study *designs* must be consistent with “scientific or economic theory.” This is both reasonable and appropriate. Missing from EPA’s language, however, is any requirement that *results* also be consistent. A competently performed risk assessment based on animal data that yields results inconsistent with human experience should not be presumed to be “sound.” Similarly, using the best stated-preference methods to estimate value do not yield “sound” results if they violate the Law of Demand.

EPA’s requirement that “assumptions, governing equations and mathematical descriptions” be “clearly justified” seems reasonable on its face. However, that EPA wants these matters *justified* rather than simply *explained* creates cause for concern. It is conceivable that EPA might be expecting third parties to include other, nonscientific attributes in their justification – such as, for example, evidence that “assumptions, governing equations and mathematical descriptions” are consistent with substantive Agency policies.

#### Bullet #4

Bullet #4 requires that surveys and other measurement instruments be validated. This requirement is reasonable provided that, like a number of others, it also applies to EPA-generated data. EPA should not impose any greater burden of proof on third parties than it actually imposes on itself.

#### “Applicability and Utility.”

EPA proposes to evaluate third-party information based on “[t]he extent to which the information is applicable and appropriate for the Agency’s intended use.” Whether information is “applicable” might be reasonably ascertained by outside parties, but the term “appropriate” can be used to mean anything one chooses. To assuage concerns that EPA will apply this criterion post hoc and without clear explanation, the Agency should provide as many examples as possible of situations where scientific information would *not* be appropriate. If it cannot identify such situations, then perhaps EPA should just delete this term from the definition. It is important that EPA not use legalistic language to impose substantive policy constraints on third-party information.

The other three bullets also seem reasonable, even if they do not convey much information. Language about the timely relevance of survey information is welcome. Out-of-date survey information is doubly problematic because such information also may not meet the “soundness” criterion due to increasing difficulties in obtaining statistically valid samples.

Despite having specifically mentioned surveys, EPA does not say anything about the Paperwork Reduction Act (PRA) for which there is statutory and regulatory language for “practical utility” combined with more than 20 years of interpretative history. This omission is not a serious one insofar as EPA is subject to PRA requirements but non-federal third parties are not. As a way of leveraging scarce resources, however, EPA increasingly seeks to collaborate with third parties to conduct research. It would be useful for EPA to remind third parties that when they work with EPA this way, they are in fact subject to the PRA because EPA’s collaboration renders it sponsored research.

“Clarity and Completeness.”

This assessment factor appears to be a rewrite of the transparency requirement contained in OMB’s government-wide directive. Indeed, EPA acknowledges the relationship in footnote 3, stating that

“transparency” generally refers to the clarity and completeness with which data, assumptions, and methods of analysis are documented, such that replication is possible if information is sufficiently transparent.

If this assessment factor is intended to implement OMB’s transparency requirement, why did EPA feel compelled to invent the new terms “clarity and completeness”? One reason for confusion is EPA states that *all five* assessment factors “taken together” “address the transparency of information.” Frankly, this simply makes no sense. Transparency is needed to assure reproducibility, which itself is a necessary but insufficient condition for demonstrating achievement of a high data quality standard. The other four assessment factors address substantive data quality attributes and are not mechanisms for ascertaining whether data quality attributes have been achieved.

Third-party information should be subject to exactly the same standards for transparency and reproducibility that apply to EPA-generated information. With respect to this data quality element, third-party information should prevail in every contest with EPA-generated information where third-party information is transparent and reproducible and EPA-generated information is not. The converse is also true: EPA-generated information should prevail on this data quality element where it is transparent and reproducible and third-party information is not.

“Uncertainty and variability.”

Uncertainty and variability are surely important data attributes, and EPA is right to consider them when evaluating the quality of third-party data. However, three of EPA’s four explanatory bullets do not provide much insight concerning why uncertainty and variability are important. Bullets #1 and #3 fit could just as easily have been located under the “soundness” criterion because they address matters related to scientific, statistical or laboratory methods. Moreover, the definition itself seems to focus primarily on whether uncertainty and variability are adequately characterized rather than whether they are unduly large. This seems more like a “clarity and completeness” issue. Only bullet #2 deals squarely with why uncertainty and variability matter; it is because they “impact the conclusions that can be inferred from the data and the utility of the study.” A large part of what matters here is whether third-party information is subject to less uncertainty than the EPA information which it would supplant, but that brings us back to the question of whether EPA intends to impose differential standards.

Bullet #4 contains welcome language related to measurement error. Too often, statistical analyses simply assume away all types of error except sampling error. That is, data are assumed to be fixed rather than realizations of some systematic or stochastic process that is dependent on, or possibly affected by, the mere act

of data collection. This practice overstates the confidence one ought to place in statistical results. Of course, the same can be said about EPA-generated data, so it remains unclear why EPA intends to apply these assessment factors to third-party data rather than to *all* data it uses or disseminates.

“Evaluation and review.”

This assessment factor appears to be akin to a strength-of-evidence test. Third-party information that has been peer reviewed, validated and replicated is more credible than information which has not. However true this might be, it is not a novel insight. Each of the four explanatory bullets seems reasonable, but not all of them appear to apply. Bullets #3 and #4 are aspects of “clarity and completeness” to the extent that these latter terms of art really mean “transparency.” Bullet #2 might relate to “soundness” and bullet #4 seems to speak to reproducibility – an assessment factor that may be contained within “clarity and completeness,” or may be missing altogether.

Bullet #2 says that third-party information will be evaluated based on the extent to which a “procedure, method or model” has been used elsewhere and whether results are “consistent with other relevant studies.” Neither of these considerations belongs at all. EPA should not establish a standard that requires third-party information to say the same thing as other information in EPA’s possession.

How does EPA intend to apply the proposed assessment factors? This cannot be ascertained, for all of EPA’s statements are couched in ambiguity permitting wide interpretative discretion. Leaving that aside, it is also unclear how EPA intends to weight these factors in its information quality determinations. Why should third parties have confidence that EPA will not apply case-specific post hoc rationales to reject third-party information it doesn’t like? Such a concern is not irrational, for in its proposed information quality guideline EPA stated that this is precisely what it intended to do:

The guidelines may not apply to a particular situation based on the circumstances, and EPA retains discretion to adopt approaches on a case-by-case basis that differ from the guidelines, where appropriate, gray decisions regarding a particular case, matter or action will be made based on applicable statutes, regulations and requirements. Interested parties are free to raise questions and objections regarding the substance of the guidelines and the appropriateness of using them in a particular situation. EPA will consider whether or not the guidelines are appropriate in that situation.

How do these assessment factors apply to co-regulators such as States which submit information pursuant to the requirements of federal regulatory programs? EPA acknowledges that “other federal, state, tribal, local, and international agencies” are third parties under the Data Quality Act. They are indistinguishable from other third parties -- except to the extent that they wield co-regulatory powers. Special procedural requirements ought to apply to information submitted to EPA by co-regulators. In particular, when they submit information to EPA wearing their hats as co-regulators, they are not mere third parties but extensions of EPA and should be subject to the same rules and procedures that apply to EPA.

## CONCLUDING COMMENTS

Many public commenters objected to EPA’s proposed information quality guideline because it was silent with respect to the critical issue of how the Agency would handle third-party information. It is understandable that, when faced with these comments, EPA concluded that it needed to finally resolve known internal disputes and get something in the public record before October 1<sup>st</sup>.

The larger problem is that EPA apparently has not resolved known internal disputes that prevented it from including applicable language in its proposed guidelines. This document appears to be a delicately balanced product of a committee of lawyers charged with writing enough text to be taken seriously but not so much that it actually says anything.

EPA should reconsider its entire approach to third-party information. The Agency needs to establish a small number of governing principles. I propose five:



1. Third-party and Agency-generated information are subject to the same standards for information quality when used or disseminated by EPA.
2. Procedures and requirements that by law, regulation or other authority apply to EPA as a federal agency do not apply to third parties unless they voluntarily collaborate with EPA or another federal agency to produce information that the Agency disseminates or uses.
3. States and other co-regulators who submit third-party information in their capacity as co-regulators must meet all requirements that apply to EPA before the Agency can disseminate or use this information.
4. Third-party information must be preferred in any case where it is superior in quality to EPA-generated information that the Agency would otherwise disseminate or use.
5. In no case must third parties satisfy information quality standards that EPA itself does not meet for an equivalent purpose.

With an appropriate amount of time, a somewhat longer or more carefully refined list might have been possible.

I appreciate the opportunity to comment. As always, I would be happy to respond to inquiries or help in any way.

Sincerely,



Richard B. Belzer, Ph.D.  
President