On Objective Risk

Dima Yazji Shamoun and
Edward J. Calabrese

July 2015

MERCATUS WORKING PAPER

MERCATUS CENTER
George Mason University

3434 Washington Blvd., 4th Floor, Arlington, Virginia 22201
www.mercatus.org

Dima Yazji Shamoun and Edward J. Calabrese. "On Objective Risk." Mercatus Working Paper, Mercatus Center at George Mason University, Arlington, VA, July 2015.

Abstract

Objectivity in the science of risk plays a monumental role in the projection of the benefits from health and safety regulations, which constitute the majority of total reported benefits of all federal regulations. Claims concerning the accuracy of regulatory risk assessments have been untestable so far in that they focus on whether a risk assessment over- or underestimates the risk of exposure to certain hazards; yet such claims imply that a true level of risk is known. This paper proposes moving the debate from the realm of the untestable to the realm of the testable through study of the *process objectivity* of the science of risk. Consistency in adhering to a process that is meant to produce objectivity should yield objective results. This paper consolidates the existing body of guidelines and recommendations on sound risk assessment practices produced by the federal government and by various scientific bodies. It proposes that, to test the process objectivity of the science of risk as applied by regulatory agencies, a third party chosen from outside the agencies conduct a *systematic* assessment of major regulatory risk assessments, according to consolidated principles. The proposed process is testable, is objective, and—if adhered to consistently—has the potential to shed light on the accuracy of the benefits calculus of major federal health and safety regulations.

JEL codes: I18, D61, D81

Keywords: process objectivity, health and safety regulations, benefit-cost analysis, risk assessment guidelines, independent peer review

Author Affiliation and Contact Information

Dima Yazji Shamoun
Research Fellow
Mercatus Center at George Mason University
Assistant Professor of Economics
Economics and Business Department, Southwestern University
dshamoun@mercatus.gmu.edu

Edward J. Calabrese
Professor of Toxicology
School of Public Health and Health Sciences, University of Massachusetts at Amherst
edwardc@schoolph.umass.edu

On Objective Risk

Dima Yazji Shamoun and Edward J. Calabrese

1. Introduction

Over the past four to five decades, risk-based regulations have gained significant popularity in public policy both nationally and globally.[1] Governments have begun to use risk assessments (RAs) and to require that major regulations[1] be accompanied by benefit-cost analyses.[2] Because RAs and benefit-cost analyses are linked, the accuracy of risk figures estimated by regulatory agencies is crucial in the estimation of benefits; in fact, because of the nature of risk-based regulation, a risk figure is the basis of the estimation of benefits. The Environmental Protection Agency (EPA) and the Food and Drug Administration, for example, regulate a host of potential chemical and environmental hazards, and the benefits of such regulations are specified and quantified by reductions in risk.

The use of benefit-cost analyses for health, safety, and environmental regulations comes at the recommendation of many prominent economists, including Nobel laureate Kenneth Arrow. Arrow et al.[3] emphasize this in their 1996 article in *Science*:

> Because society has limited resources to spend on regulation, benefit-cost analysis can help illuminate the trade-offs involved in making different kinds of social investments. In this regard, it seems almost irresponsible to not conduct such analyses, because they can inform decisions about how scarce resources can be put to the greatest social good. Benefit-cost analysis can also help answer the question of how much regulation is enough. (p. 221)

[1] According to Executive Order 12866:[2]

> "Significant regulatory action" means any regulatory action that is likely to result in a rule that may: (1) Have an annual effect on the economy of $100 million or more or adversely affect in a material way the economy, a sector of the economy, productivity, competition, jobs, the environment, public health or safety, or State, local, or tribal governments or communities; (2) Create a serious inconsistency or otherwise interfere with an action taken or planned by another agency; (3) Materially alter the budgetary impact of entitlements, grants, user fees, or loan programs or the rights and obligations of recipients thereof; or (4) Raise novel legal or policy issues arising out of legal mandates, the President's priorities, or the principles set forth in this Executive order.

There is always scientific uncertainty in both benefit-cost analysis and risk assessment. Treating uncertainty unscientifically in an RA (e.g., failing to conduct a thorough, and preferably quantitative, uncertainty analysis of all sources of uncertainty, including model uncertainty, or failing to characterize uncertainty altogether) can lead to major errors in the estimation of benefits. As our knowledge base continues to grow, particularly with respect to mechanisms of action, some uncertainty may be reduced, and thus risk predictions may be improved. Systematic adherence to best practices in their entirety ensures that new knowledge is vetted and incorporated into risk science when appropriate.

To highlight the important role of risk figures in the assessment of benefits, we briefly consult the Office of Management and Budget's (OMB's)[4] *2014 Draft Report to Congress on the Benefits and Costs of Federal Regulations and Unfunded Mandates on State, Local, and Tribal Entities*. This report indicates that "the largest benefits are associated with regulations that reduce risks to life" (p. 14). However, as OMB highlights, there are significant uncertainties concerning the benefits of health and safety regulations. Although one source of uncertainty in benefits relates to the different values that agencies assign to a "statistical life" (an uncertainty relating to the risk *management* decision), other sources of uncertainty relate directly to the RA itself (pp. 14–18).

According to OMB,[4] the largest percentage (63 to 82 percent) of monetized benefits reported by all regulatory agencies is due to rules issued by the EPA, and specifically 98 to 99 percent of those rules relate to the improvement of air quality—mainly reductions in particulate matter ($PM_{2.5}$) (p. 13).[2] Because of the significance of PM regulations, OMB treats the uncertainties surrounding the benefits from PM reduction with special scrutiny.

[2] It is worth noting that $PM_{2.5}$ is not a particular substance but a classification encompassing a range of atmospheric particles of diameter 2.5 μm and smaller.

As outlined in the OMB[4] report cited previously, the benefits from PM regulation are difficult to ascertain because of uncertainties about some of the key assumptions in the associated RA. These uncertainties include (a) the assumption of causation between the inhalation of PM particles and premature death (p. 15), (b) the assumed shape of the dose-response curve (which the EPA changed from a threshold to a linear assumption starting in mid-2009) (p. 16), (c) the assumption that the different species of PM are equally toxic (p. 17), and (d) the models used in projecting future emissions and air quality (p. 17). Additionally, OMB[4] stresses that even when some assumptions reflect the latest scientific and peer-reviewed research and RA tools, "inherent uncertainties in the overall enterprise must be recognized" (p. 17). If, as OMB's comment suggests is possible, an RA misestimates the risk to human health of PM exposure, the RA may possibly skew the decisions associated with PM reduction and perhaps cause vast resource misallocations.

The decision to regulate exposure to PM is just one of thousands of public policy decisions that rely on risk estimates. Should vitamin D consumption be increased or decreased?[5] How much seafood, high in beneficial omega-3 fatty acids, should people consume, given that seafood also contains mercury?[6] As for air pollution, might we even *embrace* a low dose in the form of PM for its suspected protective effects?[7] Given the scientific uncertainty, lack of scientific consensus, and political aversion to errors with unlikely but highly visible consequences,[1] there are no easy answers to questions relating to risk. Even when scientific uncertainty is very small—for example, in the number of car accidents—there are still competing choices for solutions. The multiple attributes of the decision calculus make these choices difficult as well.[8]

The debate around the accuracy of RAs has mostly centered on addressing the question of conservatism. Some scholars claim that regulatory agencies systematically overestimate risk;[9–13]

others concede conservatism, yet interpret the intentional use of conservative assumptions as part of a larger societal preference to overspend when lives are perceived to be at risk;[14] still others claim that RAs in fact *underestimate* risk.[14–18] Questions of conservatism thus far have been untestable, however, because they imply that the true level of risk is known. In other words, even if regulatory agencies have an incentive to use conservative assumptions, efforts to challenge this structure must be secondary to efforts to improve the accuracy of the overall RA process.

The question that has occupied the risk assessment arena for so long—whether regulatory RAs are accurate representations of the true level of risk in the world—is what philosophers of science call *product objectivity.* In other words, the science of risk is objective when its "products—theories, laws, experimental results, and observations—constitute an accurate representation of the external world" (Ref. 19, p. 2). Given the nature of the inquiry, the accuracy of risk figures per se is hard to measure. The complexity and uncertainty surrounding the world of risk—which results from confounders, interindividual variability, and so forth—may render an accurate risk representation impossible or at least, in even the most optimistic cases, very costly to attain.

Therefore, we propose to test instead another concept of objectivity—what philosophers of science call *process objectivity*. Under process objectivity, the science of risk is objective "to the extent that, the processes and methods [the RA] that characterize it neither depend on contingent social and ethical values, nor on the individual bias of a scientist" (Ref. 19, p. 2).

We propose to test our hypothesis that regulatory RAs are process objective by testing their consistency with federal guidelines and with recommendations of the National Research Council (NRC). Concerns about the uniformity of the approach to risk assessment across regulatory agencies and about the lack of consistency in the assessment of risk within regulatory

agencies date back to at least 1983, when the NRC[20] published *Risk Assessment in the Federal Government: Managing the Process* (the "Red Book").

Those concerns continue to this day, as evidenced by the many NRC publications promoting and stressing the importance of consistency in RAs through a list of guidelines and recommendations. Yet the question of whether there has been a systematic *in*consistency—that is, a lack of process objectivity—in the assessment of risk by regulatory agencies has remained largely untested. We believe that this silence is forced by the absence of a systematic assessment of regulatory RAs—by a source outside the regulatory agencies—according to the agencies' own metric of agreed-on guidelines (i.e., federal guidelines and NRC recommendations). Although challenges to regulatory RAs have been issued sporadically over the years, such challenges were not systematic, and their results were produced by either proponents or targets of the regulation in question.[21]

One criticism of our approach may be that product objectivity is ultimately what matters in the science of risk. Naturally, we would like to directly assess the accuracy of risk figures. Harry Collins describes this dilemma as "the experimenters' regress." More precisely,

> In order to know whether an experimental result is correct, one first needs to know whether the apparatus producing the result is reliable. But one doesn't know whether the apparatus is reliable unless one knows that it produces correct results in the first place and so on and so on *ad infinitum*. (Ref. 19, p. 6)

It is true that our proposed framework can only assess process objectivity—the apparatus producing the results—by way of testing consistency. Consistency, however, in following an unbiased procedure—one designed to produce accuracy—yields unbiased results.

Our proposed test is capable of gauging the extent to which the science of risk is objective because it involves an objective process. The process we outline in this paper is composed of two parts: the first part consists of identifying the guidelines and recommendations

of a sound RA; the second part is to have third parties, such as university-based research centers, reconduct major RAs (those that report a large reduction in risk—and thus a large presumed benefit) according to the identified set of guidelines and recommendations.[3] The two essential characteristics of the proposed process that render it objective are (a) that the guidelines and recommendations be followed in their entirety without any omission and (b) that the source of the reassessment be independent of the regulatory agency.

Recommendations and guidelines for sound RA practices released by the NRC and OMB date back to the early 1980s and provide ample material for the objective metric we seek. We do not propose our own guidelines and recommendations; instead, we consolidate the body of guidelines and recommendations already in existence into the main principles and categories that define sound scientific conduct.

In summary, we propose to apply the scientific method to test the hypothesis that regulatory agencies have consistently applied their own guidelines and methodologies to produce objective RAs. In this paper, we present the first step in such an enterprise that will form the basis of the proposed project: a list of recommendations and federal guidelines, compiled into a system of principles for objective RAs.

It is well recognized in the philosophy of science that consensus may be lacking in the short run for many reasons, including scientists' different initial attitudes and beliefs. However, "convergence theorem guarantee[s] that, as long as *novel* evidence keeps coming in, the degrees of belief of agents with very different initial attitudes will finally converge" [emphasis added] (Ref. 19, pp. 22–23). The key condition for convergence of expert opinion, therefore, is that

[3] As we will discuss further, our recommendation that assessments of RAs be conducted by parties outside the regulatory agencies is meant to reduce potential bias from conflicts of interest. Although no party can be completely without bias, certain parties may be *systematically* biased for reasons that will be discussed later.

evidence must be novel (and unsuppressed) to reduce the chances of updating one's beliefs in a biased manner. As a result, reassessing a risk according to the principles presented in this paper does not guarantee that a regulatory decision will be different. However, independent systematic reassessment of risk (i.e., the flow of novel evidence) will increase confidence that public policy decisions are consistent and that society's scarce resources will be allocated to the greatest social good. It may also point to the need to find ways to ensure agency compliance with federal and other prescriptions for RAs.

This paper is presented in four sections. Section 2 reviews major publications on sound RA practices: the Red Book,[20] *Understanding Risk: Informing Decisions in a Democratic Society*,[22] *Science and Judgment in Risk Assessment*,[23] *Science and Decisions: Advancing Risk Assessment*,[24] Memorandum for the Regulatory Working Group: Principles for Risk Analysis ("1995 Principles"),[25] and Memorandum for the Heads of Executive Departments and Agencies: Updated Principles for Risk Analysis ("2007 Principles").[26] After careful review of the recommendations and guidelines proposed in various publications of the NRC and OMB, we were able to identify major criteria necessary for sound and objective RAs, specifically four broad categories of a sound risk assessment: analysis, robustness, openness and transparency, and review. Section 3 presents our proposed system of principles for objective RAs, section 4 discusses the importance of an external peer review process, and section 5 presents concluding remarks.

2. Review of Major Documents and Recommendations on Risk Assessment

The federal government and numerous scientific publications by the NRC have produced guidelines and recommendations that, if consistently applied by regulatory agencies, would result in more objective RAs. In this section, we trace these documents and provide a brief

summary of their contributions to risk assessment. These publications constitute the sources for our proposed system of principles for objective RAs presented in section 3.

We start with *Risk Assessment in the Federal Government: Managing the Process* (the Red Book).[20] This volume was a foundational step in regulatory risk assessment, as it formed the basic structure for composing RAs by identifying four major categories of an RA—hazard identification, dose-response, exposure assessment, and risk characterization. This basic structure is still practiced today.

The Red Book[20] constitutes "a search for the institutional mechanisms that best foster a constructive partnership between science and government, mechanisms to ensure that government regulation rests on the best available scientific knowledge and to preserve the integrity of scientific data and judgments in the unavoidable collision of the contending interests that accompany most important regulatory decisions" (p. 1). Although there have since been several publications on risk assessment by the NRC, the Red Book is still heavily cited. A thorough review of its guidelines and recommendations is beyond the scope of this paper, but a brief outline of its major contributions is necessary.

In the Red Book, the NRC[20] identifies three steps necessary to improve the precision of RAs (p. iii). First, in the proposed process, it is necessary to ensure that RAs "take full advantage of the available scientific knowledge while maintaining the diverse organizational approaches to [the] administration of risk assessment needed to accommodate the varied requirements of federal regulatory programs" (p. 151). Second, the Red Book stresses the importance of developing uniform inference guidelines (or defaults) to achieve "standardization of analytic procedures among federal programs" (p. 151). Third, recognizing that the state of scientific knowledge is constantly evolving, the Red Book emphasizes the need for creating "a mechanism

that will ensure orderly, continuing review and modification of risk assessment procedures as scientific understanding of hazards improves" (p. 151).

In pursuit of these objectives, the Red Book[20] makes several recommendations, and the following are of particular significance: regulatory agencies are to maintain a clear theoretical demarcation between a *risk assessment*, which is of a scientific nature, and a *regulatory decision*, which is of a political, economic, and social nature (pp.19, 37, 48). RAs, according to the Red Book, "describe, as accurately as possible, the possible health consequences of changes in human exposure to a hazardous substance; the need for accuracy implies that the best available scientific knowledge, supplemented as necessary by assumptions that are consistent with science, will be applied" (p. 151).

It is essential, therefore, according to the Red Book,[20] for a federal agency to prepare and make accessible to the public in a timely manner a written RA prior to a regulatory decision (p. 153). The RA should clearly communicate which conclusions were justified on scientific grounds and which were justified on policy grounds (p. 153). In addition, the RA must identify the weight-of-evidence methodology used, along with the quality of the evidence relied on, and it should clearly indicate any guidelines or defaults used in interpreting the evidence (p. 153). Although defaults must be accompanied by the scientific rationale justifying their use, so must any departure. Given that scientific knowledge is never complete, the RA must clearly indicate how gaps in scientific knowledge were dealt with and must characterize any resulting uncertainties (pp. 153–154).

The Red Book[20] stresses the necessity that an RA undergo a peer review by an independent scientific advisory panel (either an already established panel that is authorized by law or one created for the specific RA at hand), whose evaluations must be made publicly

available before any major regulatory action (p. 156). The Red Book also deems it essential that a board of experts develop an inference option or default framework to be applied uniformly across regulatory agencies (pp. 166–170). Adherence to the framework is recommended—but is not binding—to ensure that it remains open to challenge and flexible to accommodate scientific developments. The development and use of (or failure to use) defaults also must be open to public scrutiny, along with which defaults and assumptions were adopted as a result of scientific consensus and which were adopted for policy reasons.

One major recommendation of the Red Book[20]—echoed in later publications—is emphasis on the importance of uncertainty characterization in decision making:

> [T]he degree of uncertainty may be masked to some extent when, in the final form of an assessment, risk is presented as a number with an associated measure of statistical significance. If they are to be most instructive to decision-makers, assessments should provide some insight into quantitative characteristics of the data and interpretations that may impute more or less certainty to the final results. (p. 165)

Although regulatory agencies promptly adopted the recommendations of the Red Book in their RA guidelines,[27–31] their practice has been inconsistent (Ref. 24, pp. 6, 8, 10, 36, 42). Consequently, the NRC and OMB have published further recommendations and guidance documents, briefly reviewed below.[22–24,26,32]

The NRC's[23] 1994 report *Science and Judgment in Risk Assessment*, ordered by Congress as part of the Clean Air Act Amendments of 1990, reiterated many recommendations found in the Red Book. *Science and Judgment* was directed at EPA because of congressional concerns about matters of uncertainty, model validation, interindividual variability, data gaps and the merits of some default assumptions, and lack of scientific consensus on underlying evidence in EPA analyses.[33] Although the report[23] is aimed at improving EPA's future assessments of risk from exposure to 189 particular air pollutants (p. 10), its recommendations are applicable to the process of risk assessment in general.

In response to congressional concerns, *Science and Judgment*[23] offers its own recommendations. Defaults and their associated rationale would be clearly identified and their merit regularly reevaluated while engaging the scientific community to ensure that the best scientific knowledge is used (p. 8). Also, the uncertainty and predictive accuracy of the models and data used would be rigorously analyzed (pp. 9, 137). An RA would incorporate state-of-the-art biological and pharmacokinetic quantitative models when applicable (p. 10), while taking account of causal mechanisms and interindividual variability in susceptibility (p. 11). Formal uncertainty analysis is required to reflect the limits of scientific knowledge at the time of an assessment and to shed light on areas of research necessary to reduce major uncertainties. Uncertainty analysis would also serve to minimize errors of overestimation and underestimation (p. 139).

In January 1995, OMB[25] issued the first set of federal principles (the "1995 Principles") to serve as guidance to policymakers for the assessment, management, and communication of environmental, health, and safety risks. The 1995 Principles recognize the fluid nature of risk science, and thus OMB recommends the principles as general guidance, flexible to incorporate scientific advances.

According to OMB's[25] 1995 Principles, RAs should use the best reasonable scientific knowledge, reflected in both qualitative and quantitative characterization of risk. The characterization should be broad enough to inform a range of policies for risk reduction. Any assumptions, defaults, or uncertainties associated with the risk at hand must be explicitly stated in RAs, along with their justification. RAs should include all relevant hazards and endpoints and must consider the full population at risk, with special attention to subpopulations with heightened susceptibility. The outlined RA methodology should be consistently applied across hazardous

agents and events. OMB's 1995 Principles also stress the importance of a professional peer review process to the integrity and credibility of an RA (p. 2).

The 1996 NRC[22] report *Understanding Risk: Informing Decisions in a Democratic Society* elaborates on earlier recommendations by identifying "standards for analysis" (p. 100). In addition to what is outlined in the Red Book, *Understanding Risk* defines a good quantitative analysis as including the following characteristics (pp. 101–102):

- Assumptions must be tested for reasonableness, and analysis for accuracy.
- Models must be tested against experimental results and observational data.
- Unnecessary assumptions should be removed before the final analysis is reported, and the report must include a sensitivity analysis to confirm that the removed assumptions do not affect the results.
- Data sources must be identified and made publicly available, along with calculations.
- The results of the analysis should be clearly discussed, along with the conclusions they can support and the uncertainties surrounding them.

Uncertainty analysis received special attention in *Understanding Risk*.[22] Uncertainty characterization—including data, models, parameters, and calculation uncertainties—must also be indicated in an RA. According to *Understanding Risk*, uncertainty analysis is critical to a good RA and has the dual role of admitting current limitations in scientific knowledge and identifying potential for improvements (p. 109). The decision being drawn from an RA must take into account the source of uncertainty as well as its magnitude (p. 116). Uncertainty is multidimensional and may involve "the physical and technical aspects of the risk, the social and economic dimensions of the risk, or political or behavioral factors that influence the evolution of the risk and associated uncertainty" (p. 116).

Furthermore, according to *Understanding Risk*,[22] uncertainty analysis is especially important when point estimates of risk are likely to include significant errors (p. 111). In such cases, the sources and magnitude of the errors can substantially affect the implication of the results. Although uncertainty analysis sheds light on various limitations in an RA, it must not give an air of definitiveness and completeness that may induce overconfidence in the results. (p. 111). Uncertainty analysis must also distinguish between simulation results and those generated by field and clinical data, as the two types of data are not equivalent (pp. 109–116).

In 2007, OMB[26] issued another set of guidelines and principles (the "2007 Principles") to serve as goals for risk-based regulatory decision making. In addition to stressing many of the aforementioned recommendations from the Red Book, *Understanding Risk*, and the 1995 Principles, OMB's 2007 Principles adopted some of the guidance emphasized by Congress in the Safe Drinking Water Act.[26] In fact, the procedural mandates outlined in the act proved significant to future publications.[22,32,34]

As outlined in the act, the 2007 Principles stressed that an RA should include the best available, peer-reviewed science[35] and be conducted according to commonly accepted principles of soundness and objectivity. Data used in an RA for the purpose of decision making should also be peer reviewed and made publicly available. Building on OMB's *Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated by Federal Agencies* (the Information Quality Guidelines), the 2007 Principles emphasized that an RA must clearly identify the population at risk; report the expected or estimated central risk for the specified population;[4] state the lower- and upper-bound ranges of the risk estimate; characterize the uncertainty inherent in the risk analysis and any potential

[4] EPA's[36] own guidelines suggest that the agency itself does not do this; the intent of EPA's chosen defaults is to ensure that risk "is not knowingly underestimated or grossly overestimated" (p. 13).

scientific information that could aid in alleviating such uncertainty; and provide an overview of the scientific literature, including all positive and negative studies (i.e., supportive and opposing), along with methodologies in the scientific literature that serve to consolidate different views when consensus is lacking.[35]

In addition, OMB's[26] 2007 Principles state that an RA should clearly express the rationale for, and scientific opinion of, default assumptions, along with providing a list of plausible alternative assumptions and their effect on the estimated magnitude of risk (p. 8). The 2007 Principles further state that an RA should include the types of hazards involved, the endpoints, and the affected populations, along with the range of scientific evidence and weighting methodology used to compute the risk. Special attention should be paid to especially sensitive or highly exposed subpopulations (p. 8). Each RA must be peer reviewed with the highest of professional standards in mind (p. 10). Finally, standards of risk evaluation must be consistent across hazards and events (p. 10).

The 2009 NRC[24] report *Science and Decisions: Advancing Risk Assessment* sets out to improve both the technical aspects of risk assessment and its utility for decision making (p. 4). The report points out that the ability to compare risk implications is made difficult by the use of "bright lines," such as reference dose and reference concentration frameworks (p. 265). Such bright lines give a risk manager the impression of safety below—and harm above—a particular dose (p. 253). As a result, bright lines do not allow for a comparison of the risk of two different hazards, the unintended consequences of regulating a particular hazard (risk-risk tradeoffs), or the cost-effectiveness of efforts to mitigate target risk *x* versus target risk *y*. To make a credible public health policy decision, a risk manager must not only understand the RA at hand but must also be able to compare the risk to others (p. 99). As a result, the NRC recommends using fewer bright

lines and more of a probabilistic approach to risk—for example, using a probabilistic distribution of the population at risk at different doses instead of uncertainty factors when possible (p. 139).

To improve the utility of RAs to risk managers, the NRC[24] report *Science and Decisions* reiterates many of the technical recommendations for RAs but with amendments. For example, the depth of an RA (uncertainty analysis, variability analysis, etc.) must be commensurate with the significance of the risk at hand (p. 120). The report also notes the necessity of unifying the dose-response framework for different endpoints, such that both cancer and non-cancer dose-response assessments follow a systematic approach that allows for the use of probabilistic distributions of harm rather than bright-line safety levels (p. 265). In addition, mode-of-action considerations, differences in population susceptibility, and the effect of background exposure must be explicitly considered in RAs (p. 146).

According to Abt et al.,[37] *Science and Decisions* recognizes that an RA is a unique product that "serves as a primary scientific basis for informing regulations that may have national and global impact" (p. 1028). It also recognizes that while RAs are used to make decisions that affect economic conditions, they do not necessarily allow for consideration of relevant economic costs, benefits, or tradeoffs (p. 1028), all of which are indispensable to adequately judge whether a decision to mitigate a target risk is, in fact, achieving the net health protection sought. To that end, the recommendations of *Science and Decisions* are of special significance to the field of risk assessment, because they develop a framework for risk assessment within which questions of science and economics can be answered.

3. System of Principles for Objective Risk Assessments

This section presents step 1 of the process of achieving objective risk—a compiled system of principles for an objective RA. Step 2 is presented in section 4, where we discuss the importance

of independent peer review for the integrity of the scientific method. In this section, we define four main categories of objective risk assessment: analysis, robustness, openness and transparency, and review. It is important to note that each principle can be traced back to one or more of the following: the Red Book,[20] *Science and Judgment in Risk Assessment*,[23] *Science and Decisions: Advancing Risk Assessment*,[24] *Understanding Risk: Informing Decisions in a Democratic Society*,[22] OMB's 1995 Principles,[25] OMB's 2007 Principles,[26] OMB's Final Information Quality Bulletin for Peer Review (Peer Review Bulletin), the Safe Drinking Water Act, OMB's Information Quality Guidelines, and *Review of the Environmental Protection Agency's Draft IRIS Assessment of Formaldehyde*.[38]

Although arguably no one can be completely disinterested in any matter that he or she has knowledge of, a scientist can still act in an objective manner by following the best practices and procedures of his or her chosen field. Because, as previously mentioned, we cannot readily attest to the accuracy of risk estimates, only the objectivity of the process rendering the risk estimate can be tested. Thus, in a reexamination of a poorly executed RA, it may be possible to find parts of the process left undone, yet it may be impossible to point to a particular principle that was not followed and prove how the omission *caused* the results to be inaccurate, because knowledge is lacking of the true level of risk. Given this limitation, a scientist can remain true to the end goal of objective risk only with adherence to the entire body of principles and rules of conduct of sound risk assessment. A consistent application of a system of principles meant to produce objectivity will yield objective results, and such consistency is the gold standard for the field of risk assessment.

Therefore, although each principle in the process outlined in this section is independently important, the most crucial step is to complete the *entire* process. Each principle should be viewed as an ingredient in a recipe: if ingredients are omitted, the dough will not rise.

Category 1: Analysis

1. Quality:
    a. The depth of analysis should be commensurate with the significance of the risk implied (Ref. 22, pp. 101–102; Ref. 24, p. 120; Ref. 26, p. 4)
    b. The RA should describe the weight-of-evidence methodology used. It should take into account the full range of the scientific literature (positive and negative studies) (Ref. 26, pp. 5,9). Ideally, the RA should follow a systematic review of evidence to avoid subjectivity and lack of transparency associated with weight-of-evidence methodology (Ref. 24, p. 265).
    c. For both cancer and non-cancer endpoints, the dose-response assessment frameworks should be unified such that both follow a systematic evaluation of background exposures and disease processes, possible vulnerable populations, and modes of action that may affect human dose-response relationships (Ref. 24, p. 9).
    d. When possible for non-cancer dose-response characterizations, reference dose or reference concentration frameworks should use a probabilistic distribution of harm rather than uncertainty factors (Ref. 24, p. 139).
2. Accuracy:
    a. Models should be clearly defined and validated by testing against experimental results and observational data (Ref. 22, p. 101).

b. The RA should define the populations at risk and report an expected or central estimate of risk for each affected population (Ref. 26, p. 8).

Category 2: Robustness

1. Uncertainty:
    a. The RA should identify and provide a quantitative analysis (or, when not feasible, qualitative analysis) for each source of uncertainty that constitutes potential challenges for the type of risk analyzed (e.g., epistemic uncertainty, model uncertainty, interindividual variability and heterogeneity, parameter uncertainty, exposure measurement and estimation errors or misclassifications, other errors in variables, generalizability to other populations or conditions, sample non-representativeness and selection biases, biases resulting from variable coding, confounding) (Ref. 24, pp.119–122).
    b. The RA should provide both lower- and upper-bound ranges on the risk estimate (Ref. 26, p. 5).
    c. The RA should consider unintended consequences of efforts to mitigate target risk—that is, risk-risk tradeoffs (the potential for uninformed risk-risk substitution) (Ref. 24, p. 90).
2. Sensitivity:
    a. The RA should explain the rationale for, and scientific opinion of, the default assumptions and provide a list of plausible alternative assumptions in the risk analysis (Ref. 20, pp. 154, 163, 166; Ref. 26, p. 7).

b. The RA should measure the sensitivity of the results to the different assumptions or model specifications in 2(a), along with their effect on the estimated magnitude of risk (Ref. 24, p. 208; Ref. 26, p. 8).

c. In high-stakes risk estimates (i.e., situations where potentially important countervailing risks or economic costs are associated with mitigation of a target risk), the RA should report alternative risk estimates if they are comparably plausible relative to the risk estimate based on the default (Ref. 24, p. 205; Ref. 26, p. 8).

Category 3: Openness and Transparency

1. Clarity:
    a. So that readers can ascertain which inference options were used, the RA should identify any applicable guidelines relied on in interpreting the evidence, along with scientific evidence supporting the necessity of any deviations (Ref. 20, p. 154).
    b. The RA should state the uncertainties encountered and conclusions drawn, if any (Ref. 22, p. 101).
2. Accessibility:
    a. The RA should explicitly state the models, assumptions, and weight-of-evidence methodology used (Ref. 20, p. 153; Ref. 26, p. 8). In cases where the RA uses a systematic review to evaluate the evidence, the systematic review also must be explicitly stated (Ref. 38, p. 134).
    b. The data and analysis should be readily available to the public and to anyone interested in checking the results (Ref. 20, p. 153; Ref. 26, p. 5).

Category 4: Review

1. Prepublication review

a. The RA should be independently reviewed for the accuracy of its assumptions, calculations, logic, results, and interpretations (i.e., the conclusions drawn from analysis) (Ref. 20, p. 156; Ref. 22, p. 102; Ref. 26, p. 10).
b. The rigor of the peer review process should be commensurate with the significance of the risk implied (Ref. 32, p. 2668).

2. Retrospective review:
   a. The RA should be clear about identifying the uncertainty inherent in the risk analysis and potential scientific improvements or information that could aid in alleviating such uncertainty (Ref. 22, p. 109; Ref. 35).
   b. The RA should identify a process for review and revision of the defaults relied on (and the resulting conclusions) as scientific understanding of the hazard improves (Ref. 20, p. 166; Ref. 24, p. 22).

4. The Importance of External Peer Review

Section 3 outlines our proposed system of principles for objective RAs—a compilation and consolidation of the existing body of guidelines and recommendations. As mentioned, although the principles themselves are meant to reduce bias in risk estimates by providing a consistent framework, they can achieve this end only if they are consistently applied and adhered to in their entirety.

The first step on the road to achieving objective RAs is to identify the process to follow. This section presents step 2 of the process: how to ensure the process is being followed consistently. Step 2 is as essential to the overarching goal of objective risk estimates as the system of principles.

Step 2 can be achieved by an external *reassessment* of risk similar to an external peer review. So that the integrity and objectivity of final assessments can be maintained, we propose that they be conducted by a third party outside the regulatory agencies themselves, such as a university-based research center that receives no funding from either the agency producing the RA or the industry subject to a particular policy outcome. We recognize that no third party can be truly neutral or unbiased and that bias cannot be eliminated but only reduced. However, our recommendation that reassessments be done outside regulatory agencies is meant to eliminate *consistent* bias. As illustrated in this section, there is both theoretical and empirical evidence to support the claim that regulatory agencies tend to exhibit consistent bias when charged with conducting their own peer reviews. We present evidence that supports the importance of independence in peer review and its relationship to step 2 of the process.

The importance of peer review in science is not disputed; it is the gold standard for scientific publication (Ref. 21, p. 5). But some scientists feel more strongly than others about the role of peer review. Some consider peer review a foundational step in the scientific method; others see it as the scientific method itself.[21] It is undisputed that peer review is a pillar of sound science. For peer review to deliver on the promise of sound science, however, some protocols must be followed. For example, the credibility of the peer reviewers' comments hinges on the reviewers' independence of the study under review. Independence ensures that there are no pecuniary interests in the study at hand or conflicts of interest that could form the basis of bias.[21]

The importance of peer review to sound regulatory and public policy decisions cannot be overstated. Properly conducted, the peer review process provides quality control of an agency's use of science, in addition to aiding the openness and transparency of regulatory decisions.[21,39] One of its more prominent beneficial effects is that a properly conducted peer

review induces an agency to provide a clear demarcation between the scientific and the policy bases of its decision.[21]

Because there is always a policy element to any regulatory decision, peer review can ensure that the agency does not overstate scientific grounds for the decision.[21] Thus, legal scholars Ruhl and Salzman[21] envision the regulatory peer review process:

> Designed wisely, regulatory peer review can help reveal how much scientific uncertainty underlies an agency decision and can thus demand that the agency *explain how the gap was filled*. This function, we argue, can lead to greater transparency in agency decision processes and greater legitimacy of agency decisions with the public, legislatures, and the courts. [emphasis in original] (p. 8)

To preserve the credibility and integrity of the peer review process, the group of experts must be selected in a transparent way. To ensure that conflicts of interest do not affect the quality of the research and science, the reviewers must be independent (Ref. 32, p. 2665). And as OMB stresses, to ensure credibility of the process of regulatory peer review, a broad definition of *independence* must be applied. It is not sufficient to choose a group of experts who simply did not contribute to the study to be assessed. The experts also must be employed neither by the regulator nor by the regulated.[21]

The necessity of an independent external regulatory peer review process is founded on both theoretical and empirical grounds.[21] From a theoretical point of view, there is reasonable evidence to believe that employees of federal agencies face pressures to overstate the scientific basis of their own agency's regulatory decisions. First, evidence indicates that agencies make decisions under tight time and resource constraints; thus, pressing decisions may be made without enough scientific support. Second, institutional theories (e.g., the theory of agency mission focus) state that zealous agencies tend to "further their statutory mission in a single-minded fashion" (Ref. 21, p. 17). The field of economics provides other supporting institutional theories, such as public choice and agency capture. According to this school of thought, an

agency tends over time to serve the special interests of regulated parties at the expense of furthering its mission to protect the public interest.[21] Finally, personal biases of agency employees may explain why agencies select which scientific data and why, once selected, the data are used in a biased or an incomplete way.[21] As Ruhl and Salzman[21] point out:

> Personal bias can also play a role. Most biologists who work for the FWS [Fish and Wildlife Service] or NOAA [National Oceanic and Atmospheric Administration], one could reasonably imagine, care personally about conserving wildlife—that is why they became wildlife biologists and have devoted their careers to working in an agency dedicated to wildlife conservation. If the neutrality of agency biologists is not to be trusted, this argument suggests—and this is clearly an underlying premise of the "sound science" movement—it is because they are agency biologists with "shared biases," not because they are simply biologists. (p. 17)

In addition to the aforementioned theoretical evidence, empirical evidence supports the contention that the integrity of the peer review process becomes compromised if the process is made internal to the agency.[21] For example, in response to wide criticism of their opaque practices concerning scientific data and findings, the Fish and Wildlife Service and the National Oceanic and Atmospheric Administration adopted a peer review policy in 1994 to "instill greater confidence in the public and the courts" (p. 29). In 2003, however, a study by the General Accounting Office revealed that the peer review process was far from independent and, in fact, was "informal and actually seemed to invite bias" (p. 29). Ruhl and Salzman[21] describe the findings as follows:

> The report noted that "[Fish and Wildlife] Service officials told us that they have not adopted a formal procedure to assess peer reviewers' independence, and the Service does not publicly disclose . . . potential conflicts or prior involvement by its peer reviewers." Although the agency guidelines explained that "[i]ndependent peer reviewers should be selected from the academic and scientific community, Tribal and other native American groups, Federal and State agencies, and the private sector," and that "those selected [should] have demonstrated expertise and specialized knowledge related to the scientific area under consideration," it was the agencies who selected their peer reviewers, reviewed the peer reviews, and reported the results of the peer reviews. The GAO [General Accounting Office] found that FWS "peer reviewers are selected at the discretion of the field office scientists responsible for developing listing and critical habitat decisions." Not surprisingly, the study noted that the people FWS chose to serve as peer reviewers usually

> agreed with the agencies' positions. Without independence of the reviewers ensured, this process and its results simply invite charges of manipulation. (p. 29)

Thus both theoretical and empirical evidence suggest that when agencies are in charge of the peer review process they tend to systematically use it to reinforce their decisions rather than as an expert check on the accuracy of the science used and the conclusions drawn. Although assessment of an RA is different from peer review, selection of the parties to conduct an assessment should follow the standards known to be necessary for an objective peer review process.

One suggestion to increase confidence that the conditions of neutrality are met is to have OMB select the university, or group of universities, to be in charge of conducting the reassessment of risk or, alternatively, to allow competition between universities and research centers with different sources of funding (e.g., private funding, funding from the National Science Foundation) for the task. In either case, such an independent third party would, in theory, have no personal stake in the outcome of the reassessment and instead would help ensure scientific merit and accuracy. We agree with the NRC[40] regarding internal versus external reviewers: when a group of experts is not selected by an agency, the experts are more likely to be "open, frank, and challenging to the status quo" and not "constrained by organizational concerns" (p. 3).

Lastly, in the spirit of the framework set forth in *Science and Decisions*,[24] the final reassessments would be a joint product of independent natural scientists (e.g., toxicologists, risk assessors, epidemiologists) and independent social scientists (e.g., economists). Such joint work is essential if reassessments are to produce or include risk-risk tradeoffs, risk-health tradeoffs, benefit analyses, and cost-effectiveness analyses.

5. Conclusion

"There is no good way to tell when there are overestimates or serious underestimates of the risk. If we could tell, we would fix the risk assessments"(Ref. 41, p. 1553). These were the words of John C. Bailar III, coauthor of the seminal paper "One-Hit Models of Carcinogenesis: Conservative or Not?"[15] Although these words echo from 1989, they resonate still with the majority of practitioners in the risk assessment field. The present paper is our humble attempt to propose a way to move the discussion away from the realm of the untestable (detecting systematic over- or underestimation of risk) into the realm of the testable (detecting systematic inconsistencies with a process meant to yield objectivity). Identifying systematic inconsistencies along with their origin and cause will pave the road to fixing risk assessments.

This paper presents the first step toward testing and systematic assessment of significant RAs: a system of principles for objective risk. A study of the major federal guidelines and other recommendation documents by the NRC provides ample data from which we were able to extract those principles.

Science is a systematized study. We propose that our compiled system of principles be used as a tool for systematic assessment of major RAs. The principles should be understood as a *process* for generating objective risk assessment, and with objectivity defined as strict adherence to the process. As Nobel laureate James M. Buchanan[42] said, "The 'order' is, itself, defined as the outcome of the *process* that generates it" [emphasis in original] (p. 1). Here the order is objective risk assessment, and RAs attain their objective characteristics from the process generating them.

We understand that our proposed test is intricate and time consuming. However, as Albert Einstein said, "As for the search for truth, I know from my own painful searching, with its

many blind alleys, how hard it is to take a reliable step, be it ever small towards the understanding of that which is truly significant" (Ref. 43, p. 18) Enhancing public health and safety is a truly significant goal. A systematic and objective evaluation of risk is the first step toward achieving that goal.

References

1. Rothstein H, Irving P, Walden T, Yearsley R. The risks of risk-based regulation: insights from the environmental policy domain. Environment International, 2006; 32(8):1056–1065.

2. President of the United States. Executive Order 12866: Regulatory planning and review. Federal Register, 1993 Oct 4; 58(190):12866.

3. Arrow KJ, Cropper ML, Eads GC, Hahn RW, Lave LB, Noll RG, Portney PR, Russell M, Schmalensee R, Smith VK, Stavins RN. Is there a role for benefit-cost analysis in environmental, health, and safety regulation? Science, 1996 Apr 12; 272(5259):221–222.

4. OMB/Office of Information and Regulatory Affairs. Draft report to Congress on the benefits and costs of federal regulation and unfunded mandates on state, local, and tribal entities. Office of Management and Budget, 2014.

5. Reinhold V. Critique of the considerations for establishing the tolerable upper intake level for vitamin D: critical need for revision upwards. Journal of Nutrition, 2006 Apr; 136(4):1117–1122.

6. Williams R. Public interest comment on FDA's report of quantitative risk and benefit assessment of commercial fish consumption. Mercatus Center at George Mason University, Arlington, VA; 2009 Jun 5: Docket No. FDA-2009-N-0018-138.1.

7. Cox LA. Hormesis for fine particulate matter (PM 2.5). Dose-Response, 2012; 10(2):209–218.

8. Keeney RL, Raiffa H. Decisions with multiple objectives: preferences and value tradeoffs. Cambridge, UK: Cambridge University Press, 1993.

9. Breyer S. Breaking the vicious circle: toward effective risk regulation. Cambridge, MA: Harvard University Press, 1993.

10. OMB. Regulatory program of the United States government. Office of Management and Budget, 1990. Part I: Overview. Current regulatory issues in risk assessment and risk management; pp. 13–26.

11. Perhac RM. Does risk aversion make a case for conservatism? Risk: Health, Safety, & Environment, 1996 Fall; 7(4):297–303.

12. Sax S. US Environmental Protection Agency public hearing on the proposed updates to the national air quality standards for ground-level ozone. 2015 Jan 29. The exposure and risk assessment results do not support lowering the ozone NAAQS.

13. Belzer RB. The peril and promise of risk assessment. Regulation, 1991 Fall; 14(4):40–49.

14. Finkel AM. Science and judgment in risk assessment. Washington, DC: National Academy Press, 1994. Appendix N-1: The case for "plausible conservatism" in choosing and altering defaults; pp. 601–628.

15. Bailar JC 3rd, Crouch EAC, Shaikh R, Spiegelman D. One-hit models of carcinogens: conservative or not? Risk Analysis, 1988; 8(4):485–497.

16. Finkel AM. Disconnect brain and repeat after me: "risk assessment is too conservative." Annals of the New York Academy of Sciences, 1997 Dec; 837(1):397–411.

17. EPA/Science Advisory Board. An SAB report: guidelines for cancer risk assessment. Review of the Office of Research and Development's draft guidelines for cancer risk assessment by the Environmental Health Committee. Environmental Protection Agency; 1997; EPA-SAB-EHC-91-010.

18. EPA/Science Advisory Board. An SAB report: guidelines for cancer risk assessment. Review of revised sections of the proposed guidelines for carcinogen risk assessment. Review of the draft revised cancer risk assessment guidelines. Environmental Protection Agency, 1999.

19. Reiss J, Sprenger J. Scientific objectivity. Stanford Encyclopedia of Philosophy (Fall 2014 Edition). 2014 Aug 15.

20. National Research Council. Risk assessment in the federal government: managing the process. Washington, DC: National Academy Press, 1983.

21. Ruhl JB, Salzman J. In defense of regulatory peer review. Washington University Law Review, 2006; 84(1):1–61.

22. Stern PC, Fineberg HV, editors. Understanding risk: informing decisions in a democratic society. Washington, DC: National Academy Press; 1996.

23. National Research Council. Science and judgment in risk assessment. Washington, DC: National Academy Press, 1994.

24. National Research Council. Science and decisions: advancing risk assessment. Washington, DC: National Academies Press, 2009.

25. OMB. Memorandum for regulatory working group: principles for risk analysis. Environmental Protection Agency, 1995 Jan 12.

26. OMB. Memorandum for the heads of executive departments and agencies: updated principles for risk analysis. Office of Management and Budget, 2007 Sept 19.

27. EPA. Guidelines for carcinogen risk assessment. Risk Assessment Forum. Environmental Protection Agency, 2005 Mar; EPA/630/P-03/001F.

28. EPA. Guidelines for carcinogen risk assessment. Environmental Protection Agency, 1987 Aug; EPA/600/8-87/045.

29. EPA. Guidelines for mutagenicity risk assessment. Environmental Protection Agency, 1986 Sept; EPA/630/R-98/003.

30. EPA. Risk assessment and management: framework for decision making. Environmental Protection Agency, 1984 Dec; EPA 600/9-85-002.

31. EPA. Guidelines for the health risk assessment of chemical mixtures. Environmental Protection Agency, 1986 Sept; EPA/630/R-98/002.

32. OMB. Final information quality bulletin for peer review. Federal Register, 2005 Jan 14; 70(10):2664–2677.

33. Upton AC. Editorial: Science and judgment in risk assessment: Needs and opportunities. Environmental Health Perspectives, 1994 Nov; 102(11):908–909.

34. Information Quality Act of 2001, Pub. L. 106-554, Sect. 515 (2000).

35. National Drinking Water Regulations. 42 U.S.C. § 300g.

36. EPA/Office of the Science Advisor. Risk assessment practices and principles. Environmental Protection Agency, 2004 Mar; EPA/100/B-04/001.

37. Abt E, Rodricks JV, Levy JI, Zeise L, Burke TA. Science and decisions: advancing risk assessment. Risk Analysis, 2010 Apr 8; 30(7):1028–1036.

38. National Research Council. Review of the Environmental Protection Agency's draft IRIS assessment of formaldehyde. Washington, DC: National Academies Press, 2001.

39. National Research Council. Improving the use of the "best scientific information available" standard in fisheries management. Washington, DC: National Academies Press, 2004.

40. National Research Council. Peer review in environmental technology development programs: the Department of Energy's Office of Science and Technology. Washington, DC: National Academy Press, 1998.

41. Roberts L. Is risk assessment conservative? Science, 1989 Mar 24; 243(4898):1553.

42. Buchanan JM. Order defined in the process of its emergence. Literature of Liberty, 1982 Winter; 5(4):5–18.

43. Dukas H, Hoffman B, editors. Albert Einstein, the human side: new glimpses from his archives. Princeton (NJ): Princeton University Press; 1981.