9 Scientization and the Reform of the Risk Assessment–Risk Management Framework

In the early 1990s, around ten years after William Ruckelshaus officially embraced the notions of risk assessment and risk management to demonstrate how the EPA administered environmental problems, and as that new discipline materialized in the way the agency made decisions, calls multiplied to go beyond what the authors of RAFG supposedly prescribed. Having become the reference for how the agency works, the risk assessment–risk management framework was now accused of instituting a too-strict separation of science and policy. What was seen as an improvement in the early 1980s—a clearer definition of the actual roles of toxicologists, biostatisticians, exposure scientists, but also economists, policy analysts, lawyers, and other administrative officials in the formation of a decision—was now becoming part of the problem. People had embraced the roles too ardently and were enjoying the comfort and security of these delimited responsibilities too much. There was apparently not enough communication between the two. Risk decisions were produced following fixed science policies or so-called default assumptions rather than adjusting to the specifics of the data, uncertainties, and objectives of risk managers, in a constant dialogue and characterization of the risk.

This criticism was in part misdirected because RAFG was not actually calling for a separation of science and policy and the risk assessment–risk management framework had not been applied in exactly this manner in the agency. Among other things, it helped invent new modes of exchange between the ORD's scientists and regulatory programs, for instance. As the previous chapters have shown, it legitimized more, not less involvement of the managers of program offices and of the agency in the rules of risk calculation.

What this criticism reflected was the emergence of a new series of attacks on the EPA's scientific capacities and conservative policies that it applied to chemical risks. The early 1990s was a moment in which opponents of the agency chose to depict risk assessment as being strictly about science and delegitimized the application of default assumptions and science policies to risk decision-making. Industry think tanks and lobbyists, the OMB, and an aggressive Republican vice presidency, as well as scientists pushing for the application of the latest uncertainty-reducing methods, coalesced to argue that the agency was applying rigid, conservative assumptions and not using the best possible, most advanced science. This campaign compounded with the difficulty of bringing legal procedures involving major chemicals to a close. Pressure from Congress on the agency did the rest. This context of the early 1990s is one in which a different design—a scientistic, predictive one—emerged, based on the belief that uncertainty can be reduced thanks to the progress of scientific knowledge, and that an agency should not be organized to make decisions unless this is the case. This dispute translated into a contentious redesign of the EPA, with various people inside the agency, as well as audiences outside, fiercely competing to revise risk assessment guidelines and define standard procedures for risk characterization. Science, once more, was the medium of the reconstruction of the material identity and expertise of a contested agency.

Iconic Risk Assessments and Emerging Agency Failure

The stint of Bill Reilly as EPA administrator witnessed an accumulation of complicated cases of chemical regulation. In at least three major cases affecting a variety of regulatory programs run by the agency—namely, controlling dioxin, asbestos, and arsenic—external challenges in court, led by industry or environmental groups in succession and compounded by internal discussions and inconsistencies, created nearly inextinguishable regulatory controversies and corroded the credibility of the EPA's apparent formal way of making integrated decisions.

Dioxinlike compounds form a wide range of substances, sometimes described as the most potent known carcinogens based on animal studies. Epidemiological studies are not convincing, given the difficulty of estimating the levels to which people are exposed. The EPA regulates dioxinlike compounds under various statutes as a pesticide, as a contaminant found at

Superfund sites, as an air pollutant, and as a water pollutant. Since 1983, the regulation of the levels of dioxin in effluents had been the object of several episodes of tension, notably in 1983, when it was established that political appointees in the agency were unwisely trying to save the substance from regulatory intervention. The CAG had produced the first health assessment for use across the agency in 1981. In 1982, the EPA promulgated effluent limitations and technology-based standards for the pulp, paper, and paperboard industry.

For at least the next five years, the dioxin issue was low on the agency's research agenda, going from assessment to reassessment as new experimental and monitoring studies were published, and as various industries, the OMB, and environmental groups maintained pressure on the EPA to publish more or less conservative estimates of its carcinogenic potential. Following a front-page article in the *New York Times* in 1987 about the presence of high levels of dioxin in paper products, the agency formed a group to produce an integrated assessment. Meanwhile, a group in the RAF was asked to revise the earlier estimate. Departing from its traditional approach of using the upper bound, the RAF group decided to choose a risk value near the midpoint of the dose-response curve in order to increase the acceptable dose slightly.[1] This estimate, of course, did not end the process.

The agency was challenged in court by environmental groups and was under pressure to agree with the NGOs, given revelations in the press that it had cooperated with the pulp and paper industry to produce monitoring data (Powell 1999). The consent decree signed with the Environmental Defense Fund and National Wildlife Fund (NWF) included the obligation for the agency to perform a comprehensive risk assessment of sludge, water effluents, and products made from pulp in the bleaching pulp mills, deliver a multimedia risk assessment by April 1990, and then to propose regulations to control pulp sludge disposal and address the discharge of dioxins and furans into surface waters by October 1993.

Opponents to the EPA's conservatism, for their part, put pressure on the agency via a carefully staged scientific conference. The so-called Banbury conference was organized in 1990 by industry groups and by Michael Gallo, a toxicologist who had consistently battled for the consideration of models of biological mechanisms in risk assessment (or what he called "biologically based dose response models" or BBDR modeling), and against the standard and protective mode of linear extrapolation followed by the EPA since the

end of the 1970s. The aim of the conference was to discuss the possibility of a molecular analysis of the biological steps preceding the observed effects of dioxin. For Gallo, a conclusion of the conference was that "if we can't depart from the linear model in the case of dioxin, for which we know so much, then we can't do it for anything" (cited in Roberts 1991). Inside the conference, debates were intense, but by virtue of a vigorous public outreach program led by the Chlorine Institute, it appeared to the rest of the world to have resulted in a consensus that the risks of dioxin probably had been overestimated due to the conservatism of the EPA's assumptions (McGarity and Wagner 2008).

The Banbury conference left the EPA enthusiastic about the possibility of finding an actual scientific consensus to develop regulations for dioxin (Roberts 1991). In 1991, a National Institute for Occupational Safety and Health epidemiological study on 5,000 men over a 13-year period found no increase of cancer among workers exposed to low levels of dioxin. Both events pressured Administrator Bill Reilly to announce, in April 1991, a comprehensive reassessment of cancer and noncancer risks of exposure to 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) and other compounds. Scientists in the ORD and the SAB went to work on the issue once again. Meanwhile, Vernon Houk, a scientist in the Centers for Disease Control and Prevention (CDC), argued publicly that dioxin was a weak carcinogen, after all, and that the decision made in 1981 to buy land for $33 million and evacuate thousands in Times Beach, Missouri, was mistaken (Kelly 1995).

The reassessment inspired Reilly to say that dioxin was a "model risk assessment" for the agency. Two officials of the OHEA called it a "flagship risk assessment" (Preuss and Farland 1993) because the agency had made a concerted effort to put its risk assessment into discussions publicly with academic scientists, the regulated industry, and environmental groups. Yet nothing indicated that these new procedural engagements worked. Despite years of assessment and reassessment, accumulation of studies, and clarification of methods and judgments to be applied to such cases, the EPA was nowhere near closure of the case. If anything, dioxin was still a battleground (Bailey 1992). The years that had lapsed, the costs of reducing levels of dioxin in the rules that the EPA had proposed, and the millions spent studying dioxin and evaluating the resulting studies denied the exemplary nature of the reassessment process.

Unfortunately for the EPA, there were more high-profile controversies surrounding iconic chemicals in those years. Ozone was one. It had been

the object of a high-profile court case back in 1981 [*American Petroleum Inst. v. Costle,* 665 F2d 1176, 1185 (D.C. Cir. 1981)], which confirmed the agency's 0.12 ppm standard. During the 1980s, however, the problem of longer exposures to ozone surfaced: The respiratory effects of ozone appeared to last more than an hour beyond the end of exposure. In 1987 the agency's scientific committee for air toxics, CASAC, asked the agency to include new data on the effects of longer exposures. After the relevant office, the Environmental Criteria Assessment Office (a part of OHEA), had drafted a supplement to the criteria document, the committee sent a closure letter to Reilly. The CASAC was split, but some of its influential members, such as Morton Lippmann and Bernie Goldstein, argued that they had studies showing that the effects last beyond one hour of exposure. In 1989, Bill Reilly discreetly dropped the idea of revising the ozone standard. Congress was in the process of reauthorizing the Clean Air Act. As the criteria of risk and rules for regulating air pollutants were likely to change in the very short term, it seemed a bad time to try and bring to a halt the debate on the health effects of ozone that had split the CASAC.[2] In 1991, the American Lung Association sued the EPA to force the agency to complete the revision. The court ruled in favor of the association, forcing the EPA to revise the standard by March 1993.

Perhaps the case of asbestos was even more symbolic of the sort of paralysis that risk assessment seemed unable to avert, or perhaps even accentuated. Asbestos had been listed by the agency as a hazardous air pollutant as early as in 1973. Later, the OPTS envisaged regulating the substance under the TSCA. After several iterative health assessments, an agencywide point of agreement seemed to have been found in 1986, following a multimedia assessment of asbestos orchestrated by the OPTS. The office proposed banning the product, having discerned a linear, no-threshold dose-response curve, despite claims from industry groups that too much uncertainty remained for a precise shape of the curve to be determined. The risk was deemed to be at the level of 1 in 1,000 for workers, and 1 in 1 million for the general public. A ban would thus prevent around 2,650 cases of cancer every year. The rule was finalized in 1989, but with a lowered estimation of the benefits in terms of cases of cancer avoided (down to a range of 160 to 200).

Administrator Bill Reilly and John Moore, the head of the OPTS, shared the view that substantial evidence—notably a class A carcinogen classification by the IARC—was available to back the decision, despite remaining uncertainty about differences of potency between different kinds of fiber. A

symposium held at Harvard University's Energy and Environmental Policy Center, and follow-up papers in *Science*, nevertheless maintained the debate alive. It was, however, suggested that the risks linked to exposure to asbestos were lower than those related to exposure to indoor air (Powell 1999). In the ruling on *Corrosion-Proof Fittings vs. EPA* of 1990, the Fifth Circuit Court of Appeals did not uphold the ruling, arguing there was insufficient evidence to argue for these levels of risk, and that the cost-benefit analysis and cost per life saved were unreasonable. This ruling was a massive blow to the TSCA program and blocked any future attempt to ban substances under that statute (Stadler 1992; Powell 1999; Boullier 2016).

The inorganic arsenic dossier, finally, had been provisionally closed by the decision of the assistant administrator for the Pesticides and Toxic Offices—judged strange by some in the agency[3]—to allow the use of an uncertainty factor of 10, was reopened in 1989 by the combination of two events partly external to the agency (Powell 1999, 210). First, a coalition of environmental groups sued the EPA for not meeting a court-ordered deadline for promulgating this standard. Second, the SAB recommended the revision of the risk assessment performed by the RAF, in light of the likely existence of a mechanism of detoxification of the body, established by the recent methodology of physiologically based pharmaco-kinetics (PBPK)[4]—that the SAB was generally pushing the agency to embrace. An interagency working group was established to do so (concretely to revise the reference dose in the IRIS database), but it failed to agree on which uncertainty factor to apply to the experimental data. Hank Habicht, Reilly's deputy, decided to use a factor of 3 in order to arrive at an estimate that would preserve the maximum concentration limit that the Water Office had previously defined for the substance.

Although the institutional mechanisms for reviewing research and science policies had been activated on that dossier already—the SAB and the RAF notably, including the recently experimented external peer-review workshops—the issue motivated the ORD to work on the issue again separately. An ORD working group under the Health and Environmental Research Laboratory (HERL) looked at what research would solve the case, arguing that with a proper animal model, major uncertainties could be solved in just a few years. The Office of Drinking Water also asked another researcher in HERL to develop a research agenda on arsenic, particularly in the PBPK area. Other academic groups were getting involved too, such as the Society of Environmental Geochemistry and Health (Powell 1999).

The scientific battle between academic researchers, an industry-supported task force, and EPA scientists continued for years, without allowing any agreement to emerge on what research agenda to pursue, or even on whether further research would help reduce the uncertainty. The most concrete, immediate result of the process was that the EPA was unable to meet the deadline imposed by the court—determining a maximum-level concentration by the year 1995—as a result of a lawsuit initiated by the NGO Coalition.

The Postconservatism Campaign

What made those cases particularly difficult to handle was that the bureaucratic knowledge and standard designs used by the agency since the 1980s were now becoming objects of controversy. These cases were controversial because they were dealt with in an atmosphere of conflict concerning the EPA's normal method for deciding chemical risks. Terry Yosie, formerly the staff director for the SAB of the agency under Ruckelshaus, called this moment the era of "postconservatism."[5] In the postconservative era, decisionistic design, resting on the use of default assumptions to bridge risk data and risk management, started to be heavily criticized. Postconservatism is not a spontaneous movement, but rather the result of a new political configuration involving several political factors.

The first factor was the changing environmental policy platform of the George H. W. Bush administration. During the first half of his term (1988–1990), Bush conceded a few environmental policy initiatives, notably the reauthorization of the Clean Air Act.[6] But this "modest rebirth of federal regulatory programs … receded during the administration's final two years" (McGarity 1998, 8). No other major legislative initiatives were undertaken until the end of the mandate. Vice President Dan Quayle, at the head of the newly formed Competitiveness Council, and Richard G. Darman, the new head of the OMB, became much more aggressive in their supervision of the EPA. For instance, the OMB reviewed many major rules, sometimes creating long delays and complicated negotiations with the EPA before the agency could issue these rules. Nearly 80 percent of the rules sent to the OMB for review were suspended there for longer periods of time than had been foreseen by Executive Order 12291 in 1981 (Fiorino 1995, 72–73).

The OMB also opened a debate on risk assessment methods, waging a strong attack on the agency's preference for conservative extrapolation. In

1990, it targeted risk assessment in an aggressive report decrying conservatism (OMB 1990). The report compiled a critique of all of the analytical choices made in a typical risk assessment by the EPA. It attacked not only the issue of extrapolation from high to low doses, but also the question of exposure, and indeed all of the areas in which an assumption or inference is needed to arrive at a single risk number. The OMB articulated a critique against risk assessment that became very popular in those years: the so-called compounded conservatism (Burmaster and Harris 1993; Bogen 1994; Cullen 1994) or cascading of assumptions.[7] Its "current regulatory issue" argument went as follows: "Suppose there are ten independent steps in a risk assessment and prudence dictates assumptions that in each instance result in risk estimates two times the expected value. Such a process would yield a summary risk estimate that is more than 1,000 times higher than the most likely risk estimate" (OMB 1990, 26). The OMB also argued that that when the EPA took precautionary stances simultaneously at the levels of the choice of the most sensitive species, of the maximum tolerated dose to test in the animal, of a linear model of extrapolation to establish the risk at low doses of the chemical, of body weight versus surface area conversion, to extrapolate doses from rats to humans, of calculating excess risk based on the life-time cancer risk to the maximally exposed individual, and of a calculation of the excess risk based on an upper-bound estimate instead of a maximum likelihood estimate, it ended up producing a synthetic estimate of the risk that could be "a million times higher" than the "real" risk. This, in turn, supposedly led the EPA to focus on "trivial carcinogenic risks while failing to address substantial threats to life and health" (ibid., 14). The OMB illustrates this point by comparing the EPA's decision to ban EDB (see chapter 5) to its decision not to regulate aflatoxin, a naturally occurring substance found in peanut butter that is a million times more risky.[8]

The OMB did not have scientists at its disposal, but it did not shy away from concretely discussing the scientific foundation for each of these choices. On the more critical discussion of extrapolation, it compared the results obtained from five models to show that they could vary by two orders of magnitude. That movement was given a massive push, and the EPA was further shaken, because of the engagement in the controversy of the prominent biologist Bruce Ames, in the context of the debate on the regulation of the pesticide alar (see chapter 8). It was also accentuated when one of *Science*'s deputy editors, Philip Abelson, recorded all these

facts in an aggressive editorial. For Abelson, the case was clear: "Considerable evidence is already available that the standard EPA approach is outdated and more will be forthcoming as detailed studies of metabolic and physiological processes are made. Bruce Ames and his colleagues have produced substantial evidence that results of effects of huge doses of chemicals in rodents are often misleading.... The EPA still sets guidelines on carcinogenic risks based on the limited information available during the 1970s. The agency needs to update its regulations as new facts are discovered" (Abelson 1990, 1497).[9]

The attack of the OMB resonated with the actions of the chemical industry in Congress. The revision of the Clean Air Act gave the industry the excuse to join in the battle against the EPA's risk assessment standard methods. The 1990 amendments to the legislation presented a complicated package of variegated legal developments resulting from political compromises in Congress. The amendments introduced several new provisions, inter alia: to classify areas where air pollution levels persistently exceeded the NAAQS (nonattainment areas) and to develop plans to deal with each area; to tighten automobile and other mobile source emission standards; to require reformulated and alternative fuels in the most polluted areas; and to establish an acid rain control program with a marketable allowance scheme. But it also redesigned the EPA's core method, risk assessment, through Title III of the 1990 Amendments on hazardous air pollutants from stationary and urban area sources.[10] The discussions on this title centered on the fact that the EPA simply had been too slow in regulating carcinogens, or perhaps was deliberately avoiding developing standards, given the impossible conditions that it had to fulfill, by law, to regulate air pollutants (Graham 1985). Given the short deadlines and the recurrence of intractable risk debates on each substance, the EPA had succeeded in regulating just seven substances in around ten years (Sexton 1995).

Congress moved forward to change this, reducing the role of risk assessment in the formation of standards. According to the complicated "Technology First, Then Risk" approach defined in the Clean Air Act of 1990 (EPA 2000b), the Air Office now had to identify and apply the technologies that could reduce the emission of such pollutants from key sources ("maximum available control technology," or MACT), and then perform a risk assessment for this fraction of the risk that was not eliminated by these technologies by an ample margin of safety (set at the level of 1 in 1 million

risk of developing cancer). Congress helped make a manifest leap forward for environmental protection, compared to the many frustrated attempts in the 1980s to increase action against air toxics (Oren 1991). But in so doing, it spawned a "monster of complexity" (Quarles and Lewis 1991, v) and strongly restricted the autonomy of the agency in crafting how it would use science, and which science it would use to make decisions. Congress, in effect, acted as the primary risk manager rather than the leaders of the agency doing so (Goldstein and Carruth 2003).

In those years, the AIHC was still very actively lobbying Congress on the subject of risk assessment. Nine years after having pushed Congress to launch a critical study of regulatory agencies' handling of science through the NRC, it repeated the trick in a favorable, Republican-dominated Congress: it found a senator to push an amendment to the revised Clean Air Act, for Congress to provide funds to the NRC to review the EPA's risk assessment methodology for air pollutants. This manifested itself in section 112 of the amended Act, on hazardous air pollutants, in the form of a requirement for the administrator to "enter into appropriate arrangements with the NAS to conduct a review of (A) risk assessment methodology used by the Environmental Protection Agency to determine the carcinogenic risk associated with exposure to hazardous air pollutants … and (B) improvements in such methodology."[11] The section went into the details of what should be discussed: the method for assessing the carcinogenic potency of pollutants (dose-response); and the methods for calculating exposure, "for hypothetical and actual maximally exposed individuals as well as other exposed individuals."[12] The Act also required the EPA administrator to revise the guidelines for carcinogenic risk assessment, explaining whether and how the separate recommendations of the NRC and the SAB were taken into account in the revision.

The Red Book Problem

Using the NRC framework helped both to create a public representation of what was going on inside the agency—claiming a separation of risk assessment and risk management, or science and policy as in Ruckelshaus's public discourses (see chapter 5)—and to organize the interactions between scientists, regulatory analysts, lawyers, and political appointees. In the terms adopted for this study, the framework effectively functioned as a

bureaucratic screen, but the screen was no longer effective. A set of controversies had revealed new inconsistencies. The risk assessment–risk management framework no longer functioned, and Ruckelshaus's public separation of science and policy was no longer convincing, in light of these issues. The early 1990s sounded the demise of the RAFG, as calls rose to replace its basic catechism.

Opponents to protective environmental regulatory interventions argued that risk assessment and risk management were too separate, and in so doing, they advanced their own vision of risk assessors as irresponsible environmentalists who clad their preferred policies in the language of science. Richard Wilson (see chapter 2) and William Clark, two proponents of cost-benefit analysis and longtime opponents of the EPA's conservatism,[13] argued that risk managers contented themselves with a single-number estimate of risk, without searching for the uncertainties and assumptions incorporated in it, while risk assessors failed to pay any attention to the purpose of their assessment. They claimed that the separation of risk assessment and risk management created mutual irresponsibility: "[S]cientists have used the separation as an excuse to go off into a corner and do irrelevant research" and "Risk managers have used the separation as an excuse not to understand the science and to insist on a 'simple' statement of risk as a single number" (Wilson and Clark 1991, 191). The OMB, in its 1990 report on risk assessment at the EPA, argued along similar lines. The biggest problem, the office maintained, was that science-policy choices were not assumed as such by risk assessors and never owned up to by risk managers (OMB 1990). The industry argued that with such a separation, regulatory decisions were not science-based any more. Scientists were ignored. Uncertainties were neglected by policy managers, who could advance their preferred policies no matter what risk assessors demonstrated.

Supporters of environmental measures did not agree. But they still concurred with the idea that separation had become too strong in the agency. In the context of the increasingly tangible failure of risk assessment, and dose-response calculation in particular, to produce definitive and protective decisions, NGOs, with environmental health scientist Ellen Silbergeld at the forefront, had started to doubt and articulate strong criticism of the standard risk assessment process. At the same time, they pointed to possible alternatives (Levenstein and Wooding 1997; O'Brien 2000). At a 1991 congressional hearing on the strengths and limitations of science in policymaking,

Joe Thornton, a policy analyst for Greenpeace, vehemently argued against risk assessment: "The major real-world use of risk assessment has been to approve pollution. Regulators use it to set acceptable levels of pollution and to issue permits for polluting facilities. Highly technical risk assessments have also been used by industries to intimidate citizens working to protect their communities from pollution" (US Congress 1991; see also Kuehn 1996; Thornton 2001).[14] A wave of criticism of risk assessment emerged, making it an inherently negative and adverse technology of decision-making for environmentalists.[15]

Silbergeld, then with the Environmental Defense Fund and one-time member of the SAB, wrote in 1991 that the EPA had gone too far in separating risk assessment and risk management, creating the conditions of an "uneasy divorce" (Silbergeld 1991). She argued that due to a lack of communication between risk assessors and risk managers, no commonly agreed way of dealing with uncertainties could be defined. The solutions found were designed in a piecemeal fashion by whoever found themselves faced with the uncertainty. The resulting inconsistencies were particularly striking where decisions were delegated to states, regional offices, or program offices. As risk assessment had become highly sophisticated and less mechanical, states and regional offices were having a hard time defining risks. They had trouble knowing which office to trust when different offices assessed the same substances.

Thus, the scientization of risk assessment created a new problem of inconsistency, and ultimately of ineffectiveness. Risk assessment could not deliver on the promise of revealing what the risks were, or of enabling the imposition of protective policies—an idea that would become stronger only later, among environmental health scientists, as well as environmental groups. Attentive to the sociological and political determinants of what counts as science and objectivity, and emphasizing the importance of deliberation in the construction of socially robust knowledge, Sheila Jasanoff, a professor of science and technology studies, concurred. In a paper published by the *EPA Journal* in 1993, she explained that risk assessment was suffused with arcane policy choices: "core elements of the risk assessor's work: how to select among competing models, how to balance conflicting scientific inputs, when to revise prior assumptions, how to register and represent uncertainty, and when to hold out for more scientific information … are firmly planted in the policy domain. Too rigid a separation between risk

assessment and risk management seems in the light of this analysis to be both naive and misguided" (Jasanoff 1993b, 37).

Between these two positions, another point of view on risk assessment–risk management was expressed by Goldstein in a way that was more sympathetic to the dilemmas facing the EPA. He saw no reason to actually bring risk assessment closer to risk management or to revert to the decisions made in the aftermath of the EPA crisis of 1981–1983 and of RAFG. One simple reason for this, he argued, was that the separation never really existed,[16] and that risk assessors did what they did to be helpful, constructing as precise a view of the risks as possible given the presence of uncertainties. If there were one thing to fix, it would be to require that risk managers define more precisely what they expected from risk assessors. That was also the perspective of Donald Barnes, the staff director of the SAB. Looking back on ten years of thinking in risk assessment–risk management terms, he cautioned that the separation was not possible and advisable: "[S]ome degree of interaction between risk assessment and risk management is essential if the risk assessment answers are going to address the risk management questions" (Barnes 1993, 12).[17]

But whatever the real situation was inside the EPA, it meant that the risk assessment–risk management framework no longer worked as a bureaucratic screen, as if people had learned to decode what was really happening behind those notions. Tellingly, these new prescriptions came with more critical comments and a simplification of the content of RAFG. Most people now were arguing that the report prescribed the separation of risk assessment and risk management. Wilson and Clark (1991) said that RAFG and CORADM (wrongly described as having "considerable influence" on the EPA [ibid., 188]) articulated a paradigm of separation of risk assessment and risk management, which the EPA embraced, but with unfortunate results. Perhaps the single most direct caricature came from Sedman and Hadley, who spoke of the risk assessment–risk management schism being "consecrated" as a "universal" rule by RAFG (Sedman and Hadley 1992, 192).[18] But even Don Barnes, within the EPA, had a similar view.[19]

From this point onward, the nuances in the report were lost on most of the subsequent readers, who believed that RAFG prescribed separating science from policymaking, when the authors of the report actually were advancing a sophisticated, dichotomous definition of politics as "weighing policy alternatives" (NRC 1983, 3) (the politics in risk management) *and* as

choosing inferences in risk assessment (the politics inside risk assessment). That reduction of the report to a supposed model happened at a time when the RAFG started to be called more systematically the "Red Book," and its subtle recommendations perhaps supplanted by a simplified slogan of necessary separation of risk assessment and risk management. This expression appeared around 1990 (Wilson 1990) and became more common from 1992 on.[20] At a time when the EPA's structures, forms, and operations were being modeled and simplified, history was made by drawing a line between "then" and "now," with RAFG embodying the "then."

The fact that there was too much separation of risk assessment and risk management, then, could mean several things. Within program offices, the various roles and expertise of health and environmental scientists, policy analysts, and lawyers or regulation writers were separated, placed in different internal compartments. The Water Office was in such a situation, having created the OSTP in 1991 to host its entire scientific staff instead of having scientists sit next to offices in charge of regulatory programs. It could also mean that there was less interaction between scientists in program offices and OPPE policy analysts, as OPPE was less politically supported and interoffice examination of proposed rules (and options) were less lively.[21] In addition, the interaction between ORD risk assessors and program offices seemed perfectible. The action of the so-called regulatory oversight group within the ORD, tasked with these interactions, did not have satisfactory results (Powell 1999). Finally, regional offices and states continued to have a hard time following the assessments produced by the various offices for substances that they were interested in knowing more about in order to establish depollution measures for local contaminated sites. The internal newsletter *Risk Assessment Review*—the publication of which stopped in 1994—was not sufficient to circulate information about new risk assessments, or it gave incomplete information.

The risk assessment–risk management framework, once a harmonious model mixing protective and commensurative designs, that helped assembling the EPA—creating articulations between natural scientists and economists, between ORD and program offices, between headquarters and EPA regions, and so on—suddenly became the essence of a dysfunctional agency. Various factors compounded to cast doubt on this design. One was the observed failure to produce accepted decisions about chemical risks. Risk assessment work was iterative, not linear, and often led to complicated,

convoluted histories of assessments and reassessments, with many twists and turns toward low or high estimates. Industry groups could commission new studies to counter the risk assessment of the agency; research and scientists with various affiliations could also independently publish studies to push the EPA to go in one direction or another; and of course, industry or environmental groups challenged the risk assessments in court. Risk assessment, in short, was not conducive to accepted regulatory decisions.

Another factor was the development of new knowledge and techniques to assess the biological modes of action of chemicals that made the science policies used by the agency—the standard assumptions that it applied in the majority of its decisions—less credible. Under the pressure of more or less instrumental prescriptions to refine the science, analyze the uncertainties, explore variable levels of exposure, collect data and study plausible biological mechanisms before applying the no-threshold hypothesis, and to have risk assessments peer-reviewed before turning them into decisions, risk assessments seemed to have acquired a life of their own. Behind the formal question of the separation, the real problem was that RAFG had delegated the discussion of assumptions and sciences policies to scientists. As these were becoming more complex, less standard, it seemed important to discuss them more openly (Jasanoff 1993b).

To inscribe these new scientific developments into the EPA's standard design and way of making decisions, the chemical and petrochemical industries seized on the concept of risk characterization. Risk characterization was a part of the paradigm, a term coined in RAFG. But neither the concept nor the practice of it had really been developed since the early 1980s. Starting in 1989, though, the AIHC, an industry group that had led the campaign to deprive regulatory agencies of their scientific mission, felt that this aspect of "presentation of risk assessments," or of "risk communication to decision-makers," was a major weakness in the EPA's policy and practice (AIHC 1989). The AIHC took the lead in the organization of the first workshop in this area in 1989, with the financial support of the EPA and the Society for Risk Analysis. In the meeting, the EPA risk managers pooled their needs in terms of risk assessment, and on the basis of this, they drew up a list of very basic recommendations. These included that a risk assessment presentation had to be comprehensive and understandable; its applicability and usefulness had to be clearly stated, credible, and defensible; it had to address "contentious issues" and explicate the reasons for choosing "critical scientific assumptions"; and

the relevance of the risk assessment for the regulatory framework of interest had to be explicit (AIHC 1989; see Bier 2001 also on this).

While none of these recommendations was really innovative (many had been at least outlined in 1984 in the OPPE's guidelines), they were now much more detailed, and they singled out risk characterization as a specific step in the process, a new generic competence in which the agency had to demonstrate its proficiency.

Another workshop, organized two years later, resulted in a publication called *Improving Risk Characterization*,[22] which stressed the need to provide full descriptions of qualitative and quantitative elements and to include a "candid description of uncertainties" associated with each component of the assessment (AIHC 1992), in parallel with a report and article by researchers of the Harvard Center for Risk Analysis, including John Graham (Gray 1993; Gray et al. 1993). The AIHC eventually proposed its own guidelines (AIHC 1995).

While the documents emphasized the broad consensus on risk characterization during those years, they also advocated a particular version of risk characterization—one in which uncertainty was presented as irreducible and as an argument against regulatory intervention; one in which, then, risk characterization evolved to become a kind of gatekeeping mechanism so that an assessment of risk did not automatically translate into a regulatory measure. Beyond the more general and innocent language about the need to bridge science and decision-making, the AIHC guidance actually advanced a far more radical proposal. A slight change in vocabulary was promoted at the level of whom risk characterization was for. The AIHC spoke of "an iterative process designed to be interactive with end-users," of which there was a "diversity" (AIHC 1995, 216). The risk assessment process was no longer a pipeline spewing out estimations for an EPA regulation writer; instead, risk assessment was conceived of as producing transparent risk estimations available to stakeholders, including the industry. The risk characterization memorandum also deconstructed the mechanism of risk assessment: the practice of assembling different kinds of studies and evidence about the hazard, the dose, and the exposure in order to produce a risk estimate. Echoing the "corpuscular approach" promoted by the Supreme Court in the case of *Daubert v. Merrell Dow Pharmaceuticals* (McGarity 2003), the AIHC memo argued that each of the components of a risk assessment should be separately summarized and made transparent. Most important, the AIHC argued

that "any quantitative description of risk, exposure, potency, or other risk elements should be expressed as a range" (ibid., 217).

For the AIHC, much like the researchers of the Harvard Center for Risk Analysis, single point estimates of risk were unrealistic and false. Uncertainty and variability could not, and should not, be neglected and truncated by the imposition of a policy threshold that defines levels of intervention. The dispersion of results should be expressed as much as possible to avoid false positives and unnecessary interventions. The demand to include within risk characterization a comparison between the risk being studied and other comparable risks was part of the same rationale.

The kind of risk characterization that the industry pushed, therefore, aimed to go beyond the agency's mechanical risk assessment–risk management framework of the early 1980s, putting uncertainty to the fore and installing a design that would not generate regulatory measures as fluidly as the RAFG scheme seemed to do. Renewed political opposition to the EPA in Congress helped further formulate and advance this alternative bureaucratic design.

Congressional and Industry Pressure on the Passive Smoking Issue

The issue of passive smoking, or environmental tobacco smoke (ETS), provided the opportunity for the EPA's adversaries to launch intense, concentrated attacks on the agency and, indirectly, to reinvent its typical decision-making processes.

Passive smoking was one of the few prominent issues that a newly established Office for Indoor Air (inside the Air Office) was taking care of. Indeed, the reauthorization of the Superfund legislation in 1986 had mandated the EPA to look into problems of indoor air risks. And though no regulatory powers were granted to the EPA on this particular topic, it still had a role: to inform the public and deliver advice about the private management of these risks. Along with radon (discussed in chapter 8), passive smoking was an issue that the new office championed. The case seemed clear and ripe for intervention against so-called secondhand smoking: Evidence of the link between cancer and nonsmokers' exposure to the smoke generated by smokers existed.

In 1989, the EPA published a fact sheet about ETS, followed by a compendium of technical information, data, and studies to back the claim made in the fact sheet to the effect that passive smoking was responsible for approximately 3,000 lung cancer deaths annually in nonsmoking adults.

Building on previous reports by the NAS and the Surgeon General, and based on its own 1986 risk assessment guidelines, the EPA proposed that ETS be categorized as a class A carcinogen. In June 1990, the office published a policy guide indicating possible options for reducing the risks of cancer from passive smoking—namely, ways of instituting smoking bans in the workplace. The EPA started to be accused of launching an irrational crusade on secondhand smoking, but it was simultaneously accused of cozying up with the tobacco industry on this issue (Anonymous 1990). The latter acted first, strengthening the first representation of the action of the agency publicly, through the Congress.

The EPA's actions, of course, were important to the tobacco industry, but they also piqued the interest of Thomas J. Bliley, Jr., a Republican congressman from Virginia, first elected in 1980, who was the ranking minority member of the Subcommittee on Oversight and Investigations of the Committee on Energy and Commerce. Bliley contacted the tobacco company Philip Morris as early as 1983 to coordinate their action against the so-called war on tobacco policy (Muggli et al. 2004). Philip Morris extensively supported Bliley's campaign in Congress against the EPA's actions. Their first coordinated acts were to prevent the draft risk assessment of the Air Office from going through the SAB, and to compel some scientists in the dedicated SAB group to soften their position, to the benefit of the tobacco industry. This did not prevent the report from going to the SAB, and the SAB to advise Administrator Reilly to classify passive smoking as a class A carcinogen.

The first hearing took place in April 1991 before the Subcommittee on Health and Environment of the Committee on Energy and Commerce (the "Waxman hearing" on indoor air pollution), with Representatives Henry Waxman and Bliley submitting the deputy administrator of the EPA, Hank Habicht, to intense questioning. Waxman, a Democrat, queried Habicht on the true desire of the agency to take on the problem, when its public information tools seemed to have little effect, as well as the fact that the instrument of the TSCA had hardly been used in the past decade. Bliley, in turn, attacked the credibility of the agency's risk assessment, pointing to criticisms from the SAB and suggesting that a member of the SAB panel had deliberately withheld a study from consideration—one that would have decisively changed the face of the final meta-analysis. He also demanded to know from Habicht why the technical compendium had not been reviewed by the SAB, a question that Habicht ducked.

Bliley continued to pressure the EPA, with the help of Philip Morris, sending Reilly no fewer than eleven letters arguing that the agency had made various fundamental procedural and substantial mistakes (e.g., delegating risk assessment to a contractor) and was in effect biased in its science. The letters criticized the EPA for not including all relevant epidemiological studies in its review, particularly negative ones. Bliley made a connection between the case of ETS and the rules governing WOE in the agency to demonstrate a more general wrongdoing. Further, he criticized the fact that the agency and the SAB had applied a 90 percent statistical confidence threshold on epidemiological studies, thus endorsing studies that would have de facto been excluded from consideration had it applied the 95 percent limit that the 1986 guidelines imposed—or so he argued.[23]

Bliley's action was coordinated with other tactical moves by the tobacco industry. The first was the sound science campaign (Samet and Burke 2001; Baba et al. 2005), a wide public relations and scientific lobbying operation devised to attack the credibility of a range of studies demonstrating the link between cancer and smoking. The campaign consisted of installing in the public mind a high standard of "sound science," against which nearly any study could be put in doubt, which effectively paralyzed agencies that wanted to use science in their decision-making (Wagner 1995). It resonated with the chemical industry's push to revise risk assessment methods and to embrace the science of modes of action, the biologically based dose response models, pharmacokinetic models, over supposedly unscientific defaults.

On the judicial front, Philip Morris and RJR Nabisco (the makers of a number of cigarettes, including Camels and Pall Malls) sued the agency, arguing that the SAB report was unscientific, arbitrary, and capricious.[24] Their argument was not only that the studies included in the meta-analysis had too dissimilar designs to be subjected to a numerical comparison of their results, but also that the SAB had applied a 90 percent statistical level of confidence in its review, as opposed to the EPA risk assessment guidelines and the 95 percent confidence level standard. The case would be the forum for an unusual, detailed attack on the substantive, scientific grounds of the EPA's decision to classify ETS as a carcinogen (McGarity 2003a). In its final ruling, the court went beyond what constituted the hard-look doctrine and the normal deference to agency science, and found that EPA disregarded information and made findings on selective information; did not disseminate significant epidemiologic information; deviated from its Risk

Assessment Guidelines; failed to disclose important findings and reasoning; and left significant questions without answers.[25]

The Deputy Administrator's Initiative on Risk Characterization

At the time of the ETS affair and the OMB's controversial initiatives against the EPA's risk assessment methods, Habicht was also leading the federal Interagency Working Group on Risk Assessment, under the aegis of President Bush's science advisor, David Allan Bromley. The group—yet another interagency initiative after the IRLG (see chapter 2) and Ruckelshaus's IRMC (chapter 5), this time under greater supervision from the White House—was supposed to examine opportunities for collaboration on methods, research, and other issues of interest to agencies engaged in risk assessment. Habicht did not seem to push for strong changes at the interagency level. He defined no calendar and no outcome for the group. But he defended his work at the EPA, outlining the novelty and benefits of what was being experimented with in the agency: increasing peer involvement, for the reassessment of dioxin; emphasis on risk characterization, to ensure that risk assessments of the same substance in air and water (dioxin, but also arsenic) were concordant; revision of various guidelines; and more research on uncertainty analysis, pharmacokinetic models, and dose-response models.

At a June 1991 open meeting, in the presence of industry representatives, Habicht claimed that "science should not anticipate the ultimate policy decision," that "exposure assessment is a key to good risk assessment" and that uncertainty should be reduced "as much as possible" (NAS 1991). He advanced these ideas to counter the impression that the EPA had given in the ETS affair—to wit, that risk assessment was overridden by a predetermined antismoking policy and exposure assessment had been poorly executed. Habicht's presentation at the meeting showed, at least, that he was addressing risk assessment practices in such a way as to counter criticism that the agency was facing at that time on this particular issue.

Against Habicht's apparent reluctance to put out concrete new standards and guidelines for application across the federal government, the White House decided to move forward with its agenda. The OMB, with support from Vice President Quayle, started developing a new executive order to impose principles and standards for risk assessment across all agencies, particularly targeted at the EPA as in the early 1980s under President Ronald

Reagan. The Competitiveness Council was building on the efforts of a lobbyist who was well known to the president and to the OMB in particular: OIRA's former chief, Jim Tozzi. Tozzi had left the federal government in 1983 to become a business lobbyist, helping the tobacco industry and other companies fend off environmental and health regulations. Over the years, he created a number of organizational vessels—officially, think tanks—such as Federal Focus Inc. and the Institute for Regulatory Policy, and he organized a series of forum meetings to discuss the principles of risk assessment and the use of science in policy. The latter move was spurred by the tobacco industry's crusade against the regulation of ETS by the EPA. The meetings resulted in the publication of a thick report to the Competitiveness Council, *Toward Common Measures: Recommendations for a Presidential Executive Order on Environmental Risk Assessment and Risk Management Policy,* which Vice President Quayle embraced enthusiastically (Federal Focus 1991). It suggested the adoption of a new presidential order that would outline four principles: (1) regulatory agencies must develop a list of "risk priorities"; (2) science used in risk evaluations must be separated from management decisions; (3) risk management decisions must be cost justified; and (4) the magnitude of risks must be communicated in relation to familiar risks (Borelli 1992). As Reilly's mandate was coming to an end, Habicht, grilled in Congress on the passive smoking case and feeling the pressure from the imminent White House intervention on risk assessment standards, went forward with an initiative on risk characterization.

One main reason pushed Habicht to consider this concept as a legitimate response to the controversies of the day. The risk assessment–risk management conceptual tandem had lost its effectiveness in influencing the agency's decisions and representing this order to the outside world. Inconsistency in the decisions of the agency had increased, stimulating discourse about the EPA's inability to know what decisions needed to be made. Around 1990, the agency had attempted a cluster approach to bring representatives of programs together and align their risk assessments, but that project did not go very far.[26] At about the same time, the high-level panel of experts that was reviewing the state of science in the agency (the forthcoming *Safeguarding the Future* report) had come to the conclusion, from its thirty interviews across the EPA, that the agency was bad at communicating to Congress, decision-makers, and the media about the uncertainties residing in science and how they were (or were not) handled in the resulting decisions. The experts had

made a recommendation that the science advisor be responsible for making sure that "full, documented discussions" of the scientific pros, cons, and uncertainties, as decisions and policies are developed (EPA 1992e).

Another major internal problem of agency integration came from the increasing sophistication of the models and the supposed precision with which cancer risks were computed. These numbers hardly left any choice to the decision-maker but to regulate. Bill Reilly, the EPA administrator, had expressed the feeling that he sometimes was suppressed by a "tyranny of numbers"[27]—and confusing numbers at that. One particular problem was that the 10–6 threshold was not always interpreted in exactly in the same way across the EPA: some were assessed in terms of population risk, others in terms of individual risk.[28] In other words, as risk assessment was becoming more scientific—meaning, infused with more biology, on the modes of action of substances, and more computational work, with statistical models—risk managers had more difficulty making sense of the numbers. Ten years earlier, it had seemed sufficient that risk assessors clarified what their assumptions were, so that risk managers could interpret and work with their numbers. The assessment process was now much less mechanical. It involved long lists of choices of model parameters, case-by-case considerations of the biological modes of action of chemicals, and other complicated steps. In short, risk assessment was getting scientized, employing more sophisticated biological models and modes of calculations that regulatory officials were less well equipped to question.

All in all, risk characterization crystallized the controversies of the day surrounding how the agency should decide about risks. The outside pressure for the notion of risk characterization—notably the initiative of the AIHC on this concept—meant that the agency had to forge its own approach of the concept in order to be able to draw benefits from it. Habicht asked Peter Preuss, a scientist and experienced risk assessor, head of the National Center for Environmental Assessment in ORD (formerly OHEA), and a policy analyst to write the memo. It was circulated to the entire agency on February 26, 1992, articulating three main principles that needed to be implemented in all offices: "1) risk assessment information must be clearly presented, separate from any non-scientific risk management considerations; 2) key scientific information on data and methods and their uncertainties should be identified in the risk characterization and a statement of confidence should be included that identifies all major uncertainties along with comment on

Table 9.1
Categories of participants in the risk assessment–risk management process (adapted from EPA 1992d)

Groups generating risk assessments	Groups using risk assessments	Risk managers
• Scientists and statisticians in the ORD • OPTS and other program offices • Carcinogen Risk Assessment Verification Endeavor (CRAVE) • RfD/RfC (Reference Dose/ Reference Concentration) workgroups	• Program offices • Regional offices To develop regulations through existing databases (e.g., IRIS, ORD health assessment documents, CRAVE and RfD/RfC workgroup documents).	• Agency managers • Decision-makers

their influence on the assessment; 3) it is Agency policy to present information on the range of exposures derived from exposure scenarios and on the use of multiple risk descriptors" (EPA 1992d, 5).

The first principle read like an implicit procedural defense against accusations that the EPA was dictating to scientific risk assessors what evaluations they needed to turn out for EPA policies to go forward. Again, this was the industry's main point in the ETS affair: Risk assessment had been used to find evidence to support a long-determined policy crusade against tobacco. The report defined the EPA in the terms of the risk assessment–risk management framework, distinguishing three groups in the agency, as described in table 9.1.

The goal of this reassertion of the virtual architecture of risk analysis was meant to define responsibilities for the credible articulation of science with policy: Everyone in the agency had to be aware of the need not to mix scientific and nonscientific considerations when dealing with a risk assessment.

The memo discreetly instituted a major shift. Until then, and since the second term of Ruckelshaus as EPA administrator, it had been acknowledged that risk assessment was not a scientific exercise, but a mix of science (e.g., establishment of a dose-response relationship based on the administration of various doses of a chemical to a test animal) and science policies (assuming that carcinogenic chemicals are toxic at any dose and applying a linear extrapolation model from the actually tested doses to infer toxicity at low doses). It argued, in contrast, that risk assessment could be cleansed of

"non-scientific considerations" (EPA 1992d, 4), as if it were an exercise that could be separated from the administrative, regulatory actions of the agency. The memo was reflecting the ongoing change in the meaning of risk assessment outside the agency: its scientization, under the influence of substantial knowledge of biological mechanisms of risk causation (physiologically based pharmacokinetic modeling, especially biologically based dose response modeling). It seemed to imply that the determination of science policies had to be brought under the control of risk managers. They would no longer be set in guidelines and would not be a form of extension or codification of a risk assessor's preferred professional standpoint. They needed to become debatable scientific choices, to be made explicit to the managers of the agency.

The memo leveraged two other concepts to bring to life a new representation of the EPA's science and decision-making mode. One was uncertainty analysis. The former guidelines for risk assessment were not strong on uncertainty analysis, or even on uncertainty characterization. The carcinogen risk assessment guidelines of 1986, for example, discussed the inferences made by risk assessors but did not venture into the formalization of uncertainty analysis in any depth. They exuded a kind of professional confidence in the idea that science policies took into account the major uncertainties, and that risk assessors would be able to make their choices expertly. The guidelines noted that the results of quantitative risk estimation were uncertain, specifying that the WOE should be presented to cover this potentially important uncertainty. These guidelines did not say much more than that "major assumptions, scientific judgments, and, to the extent possible, estimates of uncertainty embodied in the assessments are presented [in the risk characterization section]" (EPA 1986a, 2).

The 1986 guidelines, in that respect, reflected a limited practical engagement with uncertainty analysis in the agency,[29] which appeared more and more unacceptable with regard to the fast developments of biological and environmental sciences on this front. In the early 1990s, uncertainty analysis in the area of environmental risk developed quite rapidly. Adam Finkel, at Resources for the Future (Finkel and Evans 1987, 1990), and Morgan et al. (1992) clarified the terms of both kinds of uncertainty to expect to see and characterize in environmental risk assessment: parameter and model uncertainties. Suter (1990) and Barnthouse (1992) made great strides exploring these methods in the context of complex ecological risk assessments. In the early 1990s, uncertainty or sensitivity analysis of biological, pharmacokinetic models appeared as well, under the impetus of biostatistician Frédéric

Bois, who showed how probabilistic methods such as Monte Carlo analysis may be used to assess and correct uncertainty in parameters (McKone and Bogen 1991; Spear et al. 1991; Woodruff et al. 1992; Woodruff and Bois 1993). John Graham's Harvard Center for Risk Analysis also addressed the topic, indicting the use of single point estimates of risk by regulatory agencies (Graham et al. 1991) and testing the elicitation of judgmental probabilities in experts to document the inevitable spread in quantitative risk estimations due to uncertainty (Evans et al. 1992).

Therefore, the Habicht memo updated the agency on these developments. It recorded the recent insistence in the scientific networks surrounding the agency on the discussion of uncertainties: "A discussion of uncertainty requires comment on such issues as the quality and quantity of available data, gaps in the data base for specific chemicals, incomplete understanding of general biological phenomena, and scientific judgments or science policy positions that were employed to bridge information gaps" (EPA 1992d, 10). Uncertainty, so the memo argued, came in the form of "measurement uncertainty"—variations in measured exposures for instance—or "data gaps." Uncertainty analysis provided qualitative characterizations of each of these two kinds of uncertainty. It went into some detail: Key scientific information about methods (e.g., choice of models and species) had to be highlighted; a statement of confidence in the assessment that identified all major uncertainties had to be provided; and information had to be presented on the range of exposures derived from exposure scenarios and on the use of multiple risk descriptors (population risk, individual risk, risk for most exposed individuals and so on). More important, it enlarged the notion of uncertainty to include the ambiguity residing in the qualitative judgments of risk assessors. In stressing the qualitative, narrative nature of risk characterization, the memo was rehearsing an old point made in RAFG[30] but often forgotten in the practice of computing numerical estimates of risk.

The second (and probably more important) policy development stemming from the Habicht guidelines was the requirement that risk assessors calculate exposure in different ways: "EPA risk assessments will be expected to address or provide descriptions of (1) individual risk to include the central tendency and high end portions of the risk distribution, (2) important subgroups of the population such as highly exposed or highly susceptible groups or individuals, if known, and (3) population risk" (ibid., 21). The guidelines did not endorse one risk assessment policy, but it translated

the outside controversy on the best way to capture risk as outlining the range of possibilities and asking its assessors to use them all, to the farthest possible extent, and justify their choices. Again, while exposure to chemicals and risks was in no way a new concept, the Habicht memo was the first occurrence of an agency-level exposure policy—at least since Alvin Alm's initiative on the issue in 1983–1984 (see chapter 7).

Of the three domains of knowledge comprising risk assessment (hazard identification, dose-response assessment, and exposure assessment), exposure assessment was the least well developed "limb" of the risk assessment process.[31] An effort was made in the ORD to develop this area (see chapter 7). A 1986 guideline on exposure assessment detailed these methods over no less than eleven pages. A subsequent document on exposure-related measurements (1988) discussed the issue even more extensively. In parallel, the Superfund guidelines, published in 1989, described in detail the range of estimates that should be presented as part of the characterization of risk, particularly in the exposure part. But guidelines did not by themselves promote exposure assessment expertise, and knowledge of the exposure of vulnerable populations continued to appear insufficient, particularly as the question of exposure of children to pesticides gained salience in the public debate at the turn of the 1990s. In 1988, while the reauthorization of FIFRA stalled in Congress, the EPA decided to bring in the NAS again, requesting a study of the particular risks of dietary pesticides to children and infants. The report came out a number of years later, in 1993.[32] *Pesticides in the Diet of Infants and Children* (NRC 1993a) "really changed the paradigm" in the Office of Pesticides.[33] It forced and guided the office to revamp its methods and sources of data for exposure work, specifically concerning the characterization of the exposure of children. It veered toward combining the distribution of food consumption and the distribution of residues, and the report led it to think about the topic of "cumulative risks" (i.e., the risks from exposure to several pesticides simultaneously) as well.[34]

Habicht had the full support of Administrator Bill Reilly in his endeavor.[35] The risk characterization initiative was a highly political one, given the pressure of the OMB and Competitiveness Council, as well as the constant oversight of Congress, not least during the ETS affair, but also the initiative of the industry on risk characterization itself. As mentioned previously, the AIHC was using this notion to argue that uncertainty should prevent the agency from deciding and that risk characterization should be the moment

of the decision-making process where assumptions of risk assessors are disclosed to and discussed with stakeholders, such as regulated businesses.

The risk characterization version of the industry was a tool to advance a very conservative agenda: minimizing regulatory intervention. Liberals and public health advocates understood the notion quite differently. For them (e.g., Finkel 1991), the whole point of risk characterization was not to produce an ever-more-refined range of estimations, which would only paralyze intervention for a longer time. Rather, it should be about determining the level of uncertainty incorporated into the calculations, in order then to make a decision either to perform further research to reduce or eliminate the uncertainty or to start making a decision and intervening to reduce the risk, the uncertainty being minimal or, on the contrary, being too large and complicated to deal with.

Habicht did not take sides for either the AIHC-Harvard version of risk characterization (as an incentive for ever-more-refined uncertainty analysis and more precise point estimates of risk) or the decisionistic version (risk characterization as a way to define levels of uncertainty and truncate scientific analysis of risks to move toward decisions). But it walked a fine line between these two versions in order to close the external debate and regain autonomy over internal reform of the agency. Like Ruckelshaus's June 1983 speech (see introduction and chapter 5), it was a moment when the internal workings of the agency were defined in a way that was presentable to a series of audiences that had conflicting views of what the agency should do. In fact, there was a strong communication initiative attached to the memo. Its immediate application in the agency was announced through a press conference, even before people inside the agency were warned of it. Reilly boasted to Congress about the initiative in 1992 hearings on the ORD's budget appropriation. The document was picked up almost immediately by various people in Congress (US Congress 1992). The memo achieved its objective—namely, discouraging Quayle from imposing any new risk assessment standards on agencies. In July 1992, the White House announced that this project had been shelved, officially because of the incompatibility of the initiative with the presidential reelection campaign.[36]

The initiative was less well designed to make an impact internally. This may have been due to its style: The content was often repetitive, and it hardly had the standardized process-and-outcome format that effective guidelines typically employ. Finkel, though very supportive (the memo

picked up on his own ideas about uncertainty), described the document as "vague" and lacking the kind of "standardization that we need to move forward" (US Congress 1992, 125). No implementation plan was attached to the memo. People felt, in any case, that they were *already* doing what the memo described (Lynch 1996, Powell 1999).[37] The Office of Solid Waste rapidly published a document on the implementation of risk characterization guidance, showing that they did not have to actually change anything to comply with the supposedly new principles (EPA 1992c).

For the time being, the risk characterization initiative helped paralyze the various prescriptions stemming from the industry and the OMB. The risk characterization concept helped to show that there was effectively a problem at the level of the interaction between scientists and decision-makers in the agency. However, at the same time, it buffered the various pressures coming from the outside in order to help the agency follow its course.

Congress and the NRC Panels as a Battlefield

Risk characterization was an end-of-mandate initiative. By the presidential election of November 1992 and the change in EPA leadership, it had become clear that it had not solved all the problems of legitimate use of science in bureaucratic action, let alone ended the controversies raging outside, among the various audiences of the agency, about the right way of linking risk assessment to risk management. In 1992 and 1993, intense discussions unfolded in Congress, during multiple hearings held at the behest of Republican representatives and senators, about the EPA's science and research and its risk assessment.[38] Assistant administrators for research and development, for toxic substances, and often the EPA administrator directly (or his or her deputy) were called to testify alongside academics standing on one side or the other, from Finkel to Graham, from environmentalists to industry representatives, to engage with the kind of science that the EPA should be producing and using, and how it should be doing so. Issues of uncertainty and exposure knowledge were recurrent in testimonies and questions raised by insistent congressmen, notably James Scheuer, Don Ritter, George Brown, Henry Waxman, and Daniel Patrick Moynihan.

Bernie Goldstein, frequently called to testify, underlined on several occasions the need to increase efforts on exposure, one of the core issues of the

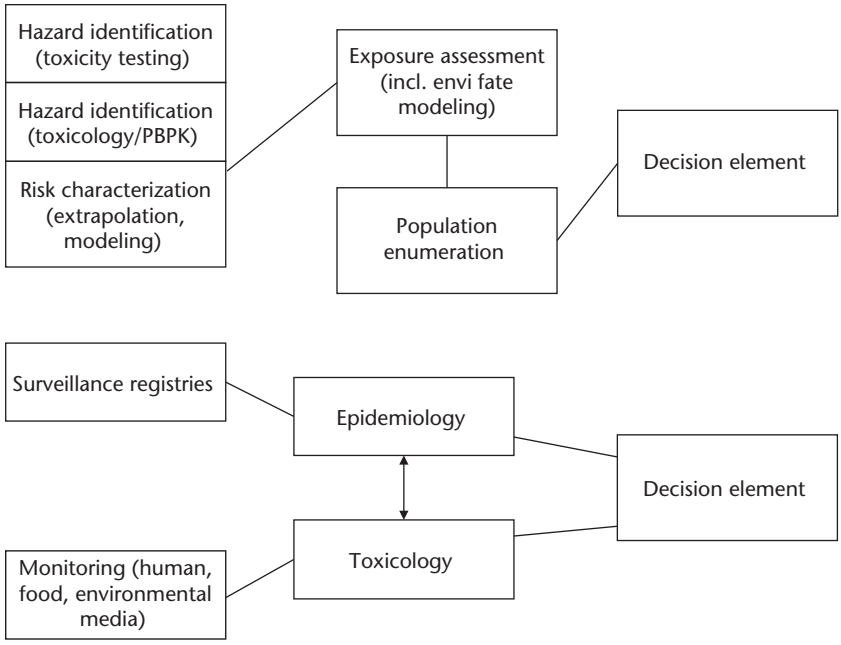

Figure 9.1
The "trinitarian" Red Book approach (top) versus a monitoring/surveillance-driven risk assessment approach (bottom) (adapted from Silbergeld 1994, 18).

risk characterization initiative. Ellen Silbergeld, then a scientist with the Environmental Defense Fund, also emphasized in Congress that epidemiological data, more than experimentally based predictions, was what now mattered more to improve the outcomes of risk assessment research, which was having such great difficulty actually delivering accuracy and protection against environmental threats. Investing in further refinements of dose-response extrapolation, through pharmacokinetics and understanding of biological mechanisms, seemed a dead end, as the case of dioxin illustrated (Silbergeld 1993). She outlined an alternative vision of what this would result in, as a kind of risk assessment process. Instead of incorporating the preference of toxicological testing and experimentally based predictions in the design (the top diagram in figure 9.1), two knowledge accumulation processes should run in parallel, both rooted in data collection and observations and united at the point of decision (bottom diagram).

Instead of simply expanding the dose-response characterization to a measure of risk for the whole population, with basic data about the

amount of substance produced and simple estimates of the doses ingested by humans, Silbergeld pleaded for constructing two separate views of the risk—one experimental, the other observational. Epidemiology, in other words, should provide independent points of reference for a policy, not simply be abducted by toxicology as exposure assessment.

These hearings did not result in any strong initiative of Congress until 1994. But in the meantime, the debate about the agency's right design for the use of science in decisions had moved to the NRC once more. In 1993, the AIHC project of the late 1970s to have the NAS review the work of regulatory agencies had resurfaced. After the publication of RAFG, managers of the NAS opposed a bill, proposed by Republican senators Don Ritter and Dave Martin, to institute a new committee inside the NAS for the purpose of reviewing the risk assessments done by agencies. However, the idea of creating a committee to work on agencies' risk assessment methodologies and guidelines persisted. The NAS ultimately refused to institutionalize a standing committee for this mission.[39] The idea nevertheless appealed to the NRC's standing expert committee that was most affected by this proposal—the Board of Toxicology.[40] The board continued to discuss it throughout the 1980s, with the resulting decision to create a dedicated committee to review the EPA's science, the Committee on Risk Assessment Methodology (CRAM).

The committee chair was given to Bernie Goldstein, the former chair of the EPA's Clean Air Scientific Advisory Committee and former assistant administrator of the ORD (1983–1986). The committee included specialists in environmental health and occupational health, as well as biostatisticians, epidemiologists, toxicologists, and ecologists. Several members of the 1981–1983 RAC sat on it, including decision scientist Warner North; Franklin Mirer, a toxicologist and industrial hygienist from United Auto Workers; and Kenny Crump, the biostatistician who had developed the mathematical model endorsed by the EPA as a basis for the agency's preferred model of extrapolation, the LMS model.

The range of expertise and viewpoints about risk assessment was wide, with supporters of the default utilization of the linear model as well as clear opponents, such as the toxicologist Michael Gallo. Three industry-affiliated scientists from Mobil Oil, Dow, and Exxon were members of the group as well. CRAM's mission was to recommend changes to risk assessment across the federal government. Its federal liaison board included representatives of

the EPA, as well as the FDA and OSHA, but its agenda was clearly focused on EPA problematics. The eventual CRAM report reinforced the emergent ideas in the fields of chemistry and biology research—namely, that risk assessment, instead of working with hypothetical statistical models to define the dose-response relationship, should make greater use of advances in biological knowledge of the modes of action of chemicals (i.e., of how substances concretely generate effects in the body at the cellular level). In other words, the report supported ongoing developments, already under consideration at the EPA, of controversial applications (see the case of dioxin discussed previously) centered around the use of pharmacokinetics and "biologically-based dose-response models," championed by Gallo. Scores of workshops and conferences took place during those years, often to present case studies revisiting the EPA's iconic regulated chemicals, to show how differently they could have been treated thanks to mechanistic knowledge (e.g., Conolly and Andersen 1993; Leung and Paustenbach 1995; Conolly 2002).[41]

The 1990 Clean Air Act amendments and the industry furthered the controversy around the risk assessment of air chemicals by the NAS. On the initiative of the industry, the amendments requested the EPA to get a report from the NRC on its risk assessment methods. Contrary to RAFG, this report had no institutional objective. The charge (which the EPA had no possibility to look into and amend)[42] requested a review of the methodologies used by the EPA, as well as of possible improvements, specifically in the areas of the assessment of carcinogenic potency and models for estimating exposure assessment. The charge explicitly pointed to the controversial "maximally exposed individual" assumption that the EPA had consistently applied since the 1970s in its risk assessments. Here, the NAS was required, directly by Congress, to evaluate the ways in which the EPA performed risk assessment.

The twenty-five-strong panel set up by the NRC in response to Congress included several biostatisticians, toxicologists, environmental and occupational health specialists, and environmental and chemical engineers. Only two lawyers were on the committee, one of whom left early in the process to join the Bill Clinton administration. A third person, Sheila Jasanoff, was also a lawyer originally, but more broadly, she was a social scientist with special knowledge in the field of science and technology studies, of which she is a leading figure. She was one of the two women sitting on the panel. There were few people there who had experience from within

regulatory agencies, except a couple of scientists who were at several points members of the SAB, and the chair of one of its three scientific advisory committees, the Clean Air Scientific Advisory Committee: Roger McClellan (who was with the Chemical Industry Institute of Toxicology when the committee convened). As is usually the case, there was no explicit trace in the report of any policy preference or explicit bias of one or several individuals on the committee. But as is usually the case as well, the NRC made sure to pick people who would represent a variety of viewpoints in order to increase the chance of forging constructive recommendations supplanting existing conflicts between any two points of view. In this case, Finkel could be singled out as the one who would carry a public health perspective to the committee, while Roger McClellan would bring into the discussion a viewpoint that reflected the chemical industry's stance on the scientific issues that were discussed. Those two scientists, along with John Bailar (a biostatistician from McGill University), John Brauman (a chemist from Stanford University), and Lincoln Moses (a biostatistician and decision scientist, also from Stanford) were said to be the most vocal people of the whole group.[43]

The charge given to the NRC via the EPA manifestly built on the grid that RAFG offered and that the EPA had since incorporated into its administrative processes and guidelines. Accordingly, one of the key members of the panel recollects that the group overall considered the framework to be sensible. It did not try to reinvent it and did not spend time on it, but on the contrary used it as a scheme to review and evaluate the EPA's activities: *Science and Judgment* (NRC 1994) showed that in the first decade of thinking in terms of risk assessment and risk management, too much separation between the two was instituted. It announced several themes that would grow in importance and would find their utmost expression in subsequent reports, including the ideas that research at the EPA should be strengthened; and that risk assessment was not an end in itself, and that its relevance for risk managers was more important than acquiring truth. A couple of years after the deputy administrator of the EPA disseminated his guidelines on risk characterization, *Science and Judgment* (NRC 1994) also emphasized the importance of that practice.

The focus of the panel's discussions was on the evolution in techniques that the EPA could develop to address the most frequent criticism leveled at it (presented on page 6 of the executive summary and further detailed

in the introduction): the use of default assumptions, the data needs, the assessment of multiple chemical exposures, and analysis and characterization of uncertainty and variability. There were, therefore, many topics on the table: a range of aspects of risk assessment that were structured according to the terms of RAFG and of EPA's existing, formal, bureaucratic knowledge. The design framed the controversy about the agency.

One topic, however, was recurrently discussed in committee meetings and prevailed in many of the panel's discussions without leading to any sort of consensus: the issue of defaults. The 1976 and 1986 guidelines of the EPA spoke of "estimations," "expert judgment," or, more positively still, of "science-policies" (EPA 1976, 1986a). By the early 1990s, however, *defaults* was the new term to designate the assumptions that the analysts of the EPA were using to calculate risks (such as the assumption that carcinogenic chemicals are toxic at any dose, and that a linear extrapolation model should be used to infer toxicity at low doses) (EPA 1990b).

Originally, the notion was used in engineering models. It then migrated to the area of exposure assessment, which was quite natural, given that this assessment did use models of dispersion and behavior in substances in various milieus. In 1992, an EPA document about pharmacokinetics applied the notion of default assumptions to the area of dose-response assessment for the first time (Federal Register 1992). Soon enough, the notion became nominalized. By then, the EPA was publicly known and decried quite simply for its use of defaults, and everyone could legitimately wonder how and when on Earth the EPA would consider moving away from these defaults, or making full use of the available science to depart from them. The default assumptions notion of environmental engineering translated into a public critique of the EPA's lack of knowledge. Mechanistic research and the developments in uncertainty and variability analysis in the area of exposure were clearly breaching the guidelines. By offering knowledge on the behavior of one of the substances in the body, the former in particular afforded an opportunity for the EPA—indeed, created pressure on the agency—to abandon its guidelines and employ the available biological knowledge rather than the standard linear extrapolation.

The theme of departure from defaults, as well as the idea that "EPA should clarify its standards for how it decides it should replace an existing default assumption with an alternative, … a point to which the entire committee agreed" (NRC 1994, 615), pervade the report. This was in spite of the fact that

the committee did not succeed in formulating a standard for when and how to depart from default options. It was precisely the issue on which the group was split. As Finkel recollects: "We all agreed on the four steps, but the issue kept coming up of defaults, assumption selection, how to bridge the gap when you have no knowledge of the fundamental chemistry … we said that the EPA had the right idea: One should have a set of assumptions, generic, reasonable, to be the starting place. For instance, positive health effects in mice are reasonable predictors. That was not controversial, including on the industry side. But in the context of the Clean Air Act, and air risk assessment, lot of questions emerged around conservatism and overregulation."[44]

These discussions turned into an open conflict which, to the great despair of NRC staff, the committee never resolved. Instead, it was decided to include two annexes in the report—one for each position in the debate for or against the use of defaults—or, as one could put it, for or against trusting the agency's expertise and reasonableness. Finkel aimed to save and support the EPA's conservatism and protective goal. He defended one position under the label of "plausible conservatism," according to which the EPA should depart from defaults whenever "there exists persuasive evidence, as reflected in a general consensus of knowledgeable scientists, that the alternative assumption (model) represents the conservative end of the spectrum of plausible assumptions (models)" (Finkel 1994, 615–616): essentially, a twofold criterion for being scientifically up to date, and consistent conservatism, both of which were designed to allow regular updates of defaults based on new validated scientific knowledge, in line with an overall policy of reducing the risks of underestimation of the risk.

The alternative view, by the decision scientist and consultant North and the toxicologist McClellan, was the "making full use of science" one. Under this approach, conservatism should not guide the departure from defaults: Abandoning a default for a more specific analysis or set of data should be motivated by the willingness to embrace the latest and most accurate science.[45] This approach was unambiguously based on the belief that uncertainty would be reduced, or even eliminated, by continuing scientific research. The hope was palpable that regulatory agencies could get rid of defaults and "unidentifiable biases" (McClellan 1994, 2003) altogether. Accordingly, they argued that Congress should do more "to encourage EPA, other federal agencies such as National Institute for Environmental Health Sciences, and private sector organizations to plan and carry out research

to reduce important uncertainties on the health consequences of toxic air contaminants" (McClellan and North 1994, 638), because in the progress of research and science lies the possibility to improve risk assessments.

In Finkel's view, uncertainty could at best be analyzed and characterized but not eliminated or reduced. Conservatism, and value choices more generally, was as inevitable as uncertainty was. In contrast, McClellan and North explicated in their text that uncertainty could and should be reduced, and that as scientific research continued, risks would be definitely and precisely known. Another significant difference between the two positions was that Finkel admitted that defining a criterion for departing from defaults, as both texts discussed, was a matter of policy. Finkel, therefore, did pronounce his policy views. He was of the opinion that the US population was favorable to a cautious heath protection policy, and that "'plausible conservatism' reflects the public's preference between errors resulting in unnecessary health risks and those resulting in unnecessary economic expenditures" (NRC 1994, 606).[46] Thus, while Finkel wrote as a scientist, with a decision theory background, as the authors of the "best science" appendix did, he also asserted clear policy choices. McClellan and North, on the other hand, did not explicitly venture onto this policy terrain. Only in the very last lines of their text did the two authors mention the "costly regulations based on conservative assumptions" (McClellan and North 1994, 640), perhaps thereby exposing their concern for the burden that health regulation creates on industries.

So, much like the RAC of 1981–1982, this committee was clearly divided into two camps. Unlike the RAC, however, the two sides were not of equal weight. Finkel admitted that, had they taken a vote, he would have been in the minority.[47] But also unlike the RAC, and unfortunately for the impact of this particular effort of the NRC, the committee did not find a way out of the conflict. Combined with the fact that the panel "tried to do too much" and "got lost in details,"[48] the conflict prevented the report from carrying a unique, neat message.

At the NRC, the report was not put forward as a major hit—quite the contrary, in fact. NRC panels are usually designed to produce recommendations that the entire panel can defend and project outside. Moreover, the NRC in general does not approve of the use of footnotes, appendices, or indeed any reflection of the dissensus expressed during committee work. Whereas RAFG and subsequent reforms in the EPA evinced an organizational order

and design that ended the controversy of the early 1980s about the agency's legitimacy as scientific bodies, *Science and Judgment* (NRC 1994) failed to have such an effect. The risk assessment–risk management framework conveyed a public image of the EPA's operations, of an agency that derives decisions from a use of science. Science-policies were an essential element of this mechanical, decision-production design. Once the agency stopped being trusted in its defaults, this mechanical design collapsed and stopped resolving controversies about the agency's legitimate action. The RAFG model was still there to provide answers. But in the context in which the fundamental expertise of the agency was put in doubt, the notion that it would now work toward ensuring a better interaction between risk assessors and risk managers to adjust assumptions case by case did not seem to reassure anyone.

Postconservatism Institutionalized

Risk characterization was the best possible design proposition that Carol Browner, EPA's administrator from January 1993 until January 2001, could use to counter attacks on the agency. She picked up on it, with a decision to update the memo of deputy administrator Hank Habicht.

Browner had originally disregarded this topic. The transition from the previous team was difficult, if not conflictual, as she wanted to launch a whole new EPA and get a divorce from previous programs and priorities in a clear-cut manner.[49] Although she claimed that she wanted to put science at the heart of the regulatory work of the agency, following one of Clinton's campaign themes,[50] her actions did not show as much interest in this resource. It took her months to be able to fill the position of head of ORD, as the general counsel of the agency refused to relax the ethical rules preventing Bailus Walker from accepting the job because he had done research under contract with the EPA. She kept William Raub in the position of science advisor created by Reilly, but she left him out of the close circle of top aides on whom she relied. Raub soon left, and the position then remained vacant for several months (Stone 1994a, 1994b). Browner, nevertheless, did act on a recommendation from the report *Safeguarding the Future* (EPA 1992e). At the end of 1993, she expanded and revamped the Risk Management Council, turning it into the Science Policy Council, tasked with reviewing and deciding science policy issues that went beyond boundaries of regional or individual programs. It took up the function that the

"oversight group"[51] led by Peter Preuss in the ORD, had, and which had been designed to coordinate the ORD's research and the science produced in program offices.[52]

Browner gradually turned to questions of science and risk analysis because she was soon faced with pressure from Republicans in Congress and their aggressive "regulatory reform" agenda deployed after the 1994 midterm elections in Congress, and the large victory of Republicans under the banner of Newt Gingrich's "Contract with America." The conservatives championed regulatory reform again and tabled a bill to mandate the agency to perform risk assessment and cost-benefit analysis for all rules.[53] Multiple proposals emerged in those years in the House or in the Senate, many of which imposed new requirements for risk analysis and cost-benefit analysis by regulatory agencies, presumably, on the side of Republicans at least, to limit the autonomy of agencies and create more occasions for review of their decisions (Graham and Sadowitz 1993; Anderson et al. 2000). In accordance with the White House initiative on risk analysis principles, Browner then resurrected risk analysis to counter regulatory reformers (Browner 1995, cited by McGarity 2001). In order not to leave this ground to Republicans, Browner decided to forge new models of operations for the agency. She relaunched, at the highest level in the agency, the interest in risk characterization. She resurrected risk characterization because the NRC argued for the need to improve this practice in *Science and Judgment* (NRC 1994). A new report, stemming from another division of the NRC, *Understanding Risk: Informing Decisions in a Democratic Society* (NRC 1996), also defended the importance of this sequence of decision-making. It was presented as a way not only to characterize uncertainties better, but also to adopt more deliberative modes of government. With the imprimatur of the Academies, the concept could help deflect and control external pressures to perform more analyses. A final reason was that Habicht's memo of virtually had no effect on the agency.[54]

Browner and her deputy set in motion a Risk Characterization Policy implementation team, largely attuned to what was in the memo of 1992. The assumption of the team was that people inside program offices were indeed separated by a "brick wall"[55] that needed to be brought down. Against the Harvard Center for Risk Analysis–AIHC coalition that pushed the subject of risk characterization between 1989 and 1992, risk characterization was reoriented to stress the need to explicate and discuss decision

rubrics of risk, cost, benefit, and the values injected into the decision much more transparently. The whole point of the new memorandum developed by the working group in direct connection with Browner and with support from Senator Barbara Boxer (a Democrat from California),[56] was less to refine the ranges of probability estimates, as conservatives would have liked to see, than to construct a procedure enabling the expression of values injected into the decision and the crystallization of the latter. The team worked to bring program and regional offices together around a table. The idea was to develop projects in each of these offices and to organize agencywide workshops to get their feedback, pushing with each of them the agenda of clarity and consistency in risk management, which stakeholders were so eager to see implemented across the agency. Once the memo was out in 1995, Browner triggered a new process of formalization of risk characterization, based less on formalisms than on principles, and aimed at changing the culture of program offices (EPA 2000a).

Next, Browner accelerated the revision of the iconic cancer risk assessment guidelines. *Science and Judgment* (NRC 1994) had a provocative effect there.[57] The revision had long been in the making and was indeed announced several times (and postponed just as often) in several workshops (EPA 1994b, 1994c), resulting in the industry and associated scientists growing weary of waiting (Anonymous 1994). In January 1994, she requested the Science Policy Council of the agency to prepare an analysis and response to the NRC, in cooperation with the ad hoc committee that the EPA had set up to review *Science and Judgment*. The council stuck to the agency's traditional RAFG-inspired design and bureaucratic screen: It considered that *Science and Judgment* gave support to the continued use of defaults by the EPA, and only needed to make improvements, clarifying their scientific and policy basis, the criteria for departure from them (NRC 1994).

Interestingly the proposed guidelines first turned to RAFG to clarify what a default—or "inference option," in the terminology adopted in that earlier report—actually was, and to defend its use. The reference to RAFG caused the EPA to reiterate: "Since there is no instance in which a set of data on an agent or exposure is complete, all risk assessments must use general knowledge and policy guidance to bridge data gaps" (EPA 1996a, 15), and to add that "some gaps in knowledge and data will doubtless continue to be encountered in assessment of even data-rich cases" (ibid., 19). Perhaps what these lines indicated most clearly was that the EPA did not really believe in

McClellan and North's scientistic opinion that science alone would reduce uncertainty as time passed and knowledge progressed. The reference to RAFG also supported the claim that "[t]he choice of an inference, as the report observed, comes from more than scientific thinking alone" (ibid., 20) and varies enormously from case to case.

In other words, the proposal implicitly argued, *Science and Judgment* was wrong to associate the EPA with a systematic conservatism. The proposal noted, furthermore, that criticism of the EPA varied (accusing the agency of being too resistant to changing its defaults or questioning the basis on which to justify its actual departures from defaults), as did the proposed criteria for when to depart from defaults. The proposal recounted the disagreement expressed in the NRC panel and in the appendices of the final report between Finkel and North and McClellan. Quite clearly, the conflict among experts on the issue legitimized the EPA's decision not to formalize any criteria to depart from defaults ["No uniform checklist will fit all cases … a checklist would likely become more a source of rote discussion than of enlightenment about the process" (ibid., 17)]—a stark illustration of why and how dissensus in reports failed the NRC. The proposal contained a framework whereby "[i]f data support a plausible alternative to the default, but no more strongly than they support the default, both the default and its alternative are carried through the assessment and characterized for the risk manager. If data support an alternative to the default as the more reasonable judgment, the data are used" (ibid., 17–18).

The main building blocks of the new approach were agreed upon in a meeting of the RAF in August 1994 (ibid.), and then a workshop with twenty-five outside scientists organized a month later to foster discussion. First, most of the scientists agreed with the change to a narrative form of hazard identification. The agency would label them as "Known/Likely," "Cannot Be Determined," or "Not Likely," and the disappearance of strict letter classes would eliminate the definitive, regulation-inducing decision of whether a substance was a carcinogen.[58] Second, the use of defaults was to be minimized and greater consideration given to knowledge of modes of action, thresholds,[59] biological models, and the like. It recommended against the use of the maximum tolerated dose. Most scientists, including industry consultants, hailed the shift. Finally, the stress on epidemiological data led the press to argue that the EPA was moving away from reliance on animal testing and tumor identification in rats.

The 1996 guidelines directly reflect the trends embedded in other guidelines—on exposure assessment notably, as well as on risk characterization. Exposure of sensitive populations was given greater importance in this text, which also included a lengthy description of the appropriate practices for risk characterization. What it did not do, however, to the displeasure of environmental groups, was to incorporate a method of looking jointly at chemicals that have cumulative effects. That new approach and concept was reserved for subsequent guidelines.

The EPA encountered stronger opposition to its choice to relax the criteria of statistical significance for exposure studies. The new guidelines increased and specified the criteria of the 1986 guidelines,[60] but they dropped the "unlikely to be due to chance" formulation. The deletion of this language gave the agency greater leeway to consider studies that applied other thresholds of confidence, such as a 90 percent confidence level, as it had attempted to do in its policy on electromagnetic fields and on ETS (Gough and Milloy 1996). In doing so, the EPA was possibly trying to restore a balance: on the one hand, it gestured toward considering the possibility of not applying its defaults, particularly as concerned the nonthresholded nature of chemicals. But it relaxed the statistical confidence criterion to include more epidemiological studies and identify carcinogens that would have been overlooked by the reintroduction of a threshold approach. The same compromise was found in Congress to revise the Food Quality Protection Act. With intense involvement by Browner, the Delaney amendment was repealed, but an extra uncertainty factor of 10 was inserted for cases of evidence of fetal and postnatal developmental toxicity, as well as when data from toxicity testing relative to children were incomplete (NRC 1993a).

As Preuss recollects, the revision was an episode in a "twenty-year battle" between the cadre of agency risk assessors and a host of people (both inside and outside the agency) who aimed to influence their preferred approach: "Way back, in 1983, we said draw a straight line. In 1986, we said draw a straight line. In the 1990s, we said, draw a straight line. Our position never changed, because we said we have no idea what happens at low dose, what the shape of the curve is at low dose. So default would be straight line. But if you had evidence, you could change that. So there was a lot of fighting, a lot of pressure to reverse that thinking."[61]

The agency broadly maintained its preferred approach in that revision, but it allowed some flexibility. Since the first version of the 1980s, the

guidelines had grown more adaptable and detailed over time. The flexibility that was introduced in the decision-making process, through the practice of considering, case by case, whether there was enough biological knowledge to not apply the chosen, default assumption, created more uncertainty for the EPA concerning the results of reviews initiated by the courts or elsewhere. It created the need to get more control over what scientists in the agency did, and risk characterization accomplished just that.

Conclusion

Throughout the 1980s, defining an uncontestable core of risk knowledge, demarcating the remaining uncertainties, and formalizing the policies applying in these zones of uncertainty have worked as a way to establish a legitimate functioning in the EPA. It did work effectively, to the point of becoming the foundation of a public identity for the agency, framing the perception of how it would make decisions for both external and internal constituencies. Charting the various forms of knowledge that are necessary to form a generically legitimate decision is the kind of design practice that was at the source of the risk assessment–risk management framework and its installation into the heart of the EPA during the 1980s. It is within this framework that the EPA could claim to apply something called *science* and depict itself as a science-based decision-making agency.

When this assemblage and the formalized bureaucratic knowledge and technologies it rests on are in turn contested or become the very object of the controversies that they were designed to avert, a new cycle of design sets in. This is what the period that is covered in this chapter illustrates. In the early 1990s, and more and more intensively since the technologies of science-based decision-making on which the agency modeled itself, the new design was gradually seen to fail. Risk assessment guidelines, with their set of knowledge and mechanisms to shape a decision, appeared insufficient in the light of the advances of scientific research on chemicals and environmental disorders. To be sure, many of these insufficiencies were manufactured. They were instrumental for the actors that wanted to cast doubt on the way that the EPA was making decisions, as well as on its supposedly conservative policies. Faced with violent criticism of its "default" models and the rise of new, sophisticated concepts of chemical modes of action, quantitative uncertainty analysis, PBPK, and the like—an overall

scientization of the exercise of risk assessment—the EPA needed to revise the way in which it was performing its assessments, as well as the knowledge used in the process. More than that, it needed to find a way to restore the mechanical link between science and decision-making, in a context in which the methods and data for the scientific assessment of risks diversified and made this exercise less mechanical than it had been in the past.

Risk characterization—the notion that legitimate and credible decisions can be forged only if risk assessors and risk managers agree on choice of models, statistical significance criterion, application of defaults, and so on—ended up providing the guide to this reform of the agency and of its formal knowledge and technologies. This did not happen without considerable debate, hesitation, and contention because, by now, the terms of the risk paradigm and risk-based decision-making had been appropriated by the multiple groups that want to gain influence over the way that EPA makes decisions—or over its decisions, anyway. And it certainly could not bring to a close the continuous scientization of risk assessment and the pressure to reduce uncertainties and predict hazards, which both risk scientists and adversaries of regulatory intervention exerted on the agency. More responses to these renewed controversies about the extent of uncertainties and the capacity of the EPA to solve them would be forged in time.