



HAL
open science

Managing technical reputation: Regulatory agencies and evidential work in risk assessment

David Demortain, Olivier Borraz

► **To cite this version:**

David Demortain, Olivier Borraz. Managing technical reputation: Regulatory agencies and evidential work in risk assessment. *Public Administration*, 2021, Special Issue: Symposia: Corporatization of public services: new perspectives and enduring challenges, 100 (2), pp.394-407. 10.1111/padm.12734 . hal-03185765

HAL Id: hal-03185765

<https://hal.science/hal-03185765>

Submitted on 20 Dec 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

David Demortain, Olivier Borraz

« Managing technical reputation: Regulatory agencies
and evidential work in risk assessment »

Public Administration

2021: 1– 14. <https://doi.org/10.1111/padm.12734>

Managing technical reputation: regulatory agencies and evidential work in risk assessment

David Demortain, Lisis (Laboratoire Interdisciplinaire Sciences, Innovations, Sociétés)

Olivier Borraz, CSO (Centre de Sociologie des Organisations)

Abstract

How can regulatory agencies with a technical or scientific mission forge and defend their reputation, when the technical and scientific content of their work is subject to countervailing influences and perceptions among a wide array of audiences? In this paper, we tackle this broad question, focusing on a particular episode of the European controversy over the regulatory control of exposure to Bisphenol A, during which the European Food Safety Authority altered the method by which it produced an assessment of the risk of BPA, responding to the regulatory controversy surrounding this substance. Building on the literature on organizational reputation and science and technology studies, we shed light on the work that regulatory agencies undertake to gain credibility in particular cognitive configurations of audiences. This perspective on the management of audiences and knowledge standards is central for the explanation of the decisions, policies and strategies of science-based agencies, and the way in which a technical reputation takes form in controversy-prone areas of regulation.

Keywords

Regulatory agencies, technical reputation, evidential work, risk assessment, bisphenol A, EFSA

How can regulatory agencies with a technical mission forge and defend their reputation, when the scientific content of their work is subject to many countervailing influences and perceptions among a wide array of audiences? How far can regulatory agencies work with and shape knowledge, in order to satisfy the various norms that their audiences apply to regulatory science?

In this paper, we tackle this broad question by focusing on a short episode of the European controversy over the regulatory control of exposure to Bisphenol A. In 2014, in front of surprised audiences, the European Food Safety Agency (EFSA) took a sharp turn from previous recommendations in which it had always considered Bisphenol A (BPA) to pose no major risk to health and decided to lower the acceptable dose for the substance. To justify this shift, EFSA invoked the use of new methods to compute risk, departing from accepted conventions for risk assessment, to which it generally expressed a strong attachment. In so doing, EFSA was responding to a variety of governments, NGOs and environmental health scientists across Europe accusing the agency of overlooking data and studies demonstrating the hazardousness of BPA – a situation which could undermine its image as an expert body.

According to some authors (Lofstedt, 2011; Rimkute, 2018), if NGOs dominate an issue, an agency will resort to a more performative or moral reputation as a public health protector in order to uphold its reputation (Maor, 2011). This reputation will translate into a particular scientific approach, namely a precautionary, hazard-based approach, according to which scientific uncertainties command, rather than delay, protective decisions. But if firms dominate the issue, then the agency will follow a risk-based, probabilistic approach, in line with a stricter technical reputation as a science-based organization following the latest standards of regulatory science (Maor, 2007). A change like the one EFSA applied to BPA in the middle of a major controversy would thus result from a change in the audiences that bear on the question at stake.

Yet, such an interest-group and bureaucratic politics perspective reduces “reputation to a binary or monotonic choice or outcome pertaining to what a public agency has or does not have” and risks losing the “richness of administrative behavior” (Carpenter and Krause, 2012, p. 31). It caricatures agencies’ behavior as switching from one body of knowledge to another in reaction to external pressures (Rimkute 2018).

In this paper, we build on organizational reputation literature (Carpenter, 2010) and the idea that an agency’s behavior depends on the dominant images the agency holds in the eyes of its audiences. The organizational reputation approach reveals “the ability of public bodies to act adaptively, strategically and opportunistically in developing good reputations as well as in maintaining and enhancing the stability of such reputation” (Maor, 2015; see also Gilad, Maor, & Ben-Nun Bloom, 2015); for instance, through communicational strategies (Gilad, Maor, & Ben-Nun Bloom, 2015). This can be extended to the standards of knowledge agencies mobilize in the production of expertise. More precisely, we base our approach on Carpenter’s “network of audiences” – which below we call configuration of audiences – considered as spaces of construction of organizational credibility. We supplement this theory with insights from science and technology studies, to foreground the work that they undertake to shape the representations of good science that prevail among their audiences. These combined approaches allow us to describe how agencies forge modes of demonstrations of risk that align audiences, which we term evidential work. Using this theoretical framework, we claim to provide an explanation to EFSA’s sudden shift.

In the remainder of the paper, we first outline our framework based on organizational reputational literature and science and technology studies. We then move on to describe food contact materials regulation, and the standard methods of regulatory science in the area. Going into the history of the risk assessment of BPA, we show that, whilst abiding to international

standards, EFSA came under fierce criticism from both NGOs and some national agencies with regards to the many remaining uncertainties around the potential health effects of BPA. EFSA subsequently gave greater weight to studies demonstrating the hazards of BPA, but simultaneously minimized exposure to the substance in an overall new mode of demonstration of the absence of risk. This realignment served to re-establish EFSA's position, while continuing to promote a risk-based approach that earned the agency its technical reputation on other issues and for other audiences. EFSA did not just react to demands of precaution-demanding audiences, but actively shaped the science in such a way as to create an alignment of its audiences that would be favorable to the preservation of its scientific reputation. In the last section, we discuss the organizational implications of evidential work, before offering conclusions for the study of expert agencies.

Knowledge standards and technical reputation in agency-audience relationships

The organizational reputation literature (Carpenter, 2001, 2010; Carpenter and Krause, 2012; Maor, 2015) establishes that the legitimacy of an agency and its capacity to deploy its decisions are determined by the complex of images and representations that its actions generate among diverse audiences (Carpenter, 2010). Organizational reputation is complex: its objects may be agency performance, or the “capacity for effectively achieving its ends and announced objectives” (Carpenter 2010, p. 46); its moral character, or action towards various ethical ends (e.g. the agency as “guarantor of public safety”: Maor, 2011); the extent to which it demonstrates “scientific accuracy, methodological prowess, and analytic capacity” (Carpenter 2010, p. 46), or expertise; and finally, its legal-procedural behavior, or the “justness of the processes by which its behavior is generated” (*ibid.* p. 47). The images of the agency on these various grounds are generated in the interactions between an agency and its audiences. They

empower or constrain the former in its legitimacy to act and control a regulatory niche (Busuioc, 2016).

Like other dimensions of reputation, expertise needs to be approached as a symbolic construct (Carpenter, 2010, p. 46), subject to competing representations of authoritative knowledge: “what legitimates an organisation is a *perception of competence* among multiple audiences ‘that matter’.” (Busuioc and Rimkute, 2019, p. 1262, emphasis in original). A reputation of expertise will necessarily evoke various “political/ideological expectations” (ibid., p. 1263) that bear on what is defined as expertise and on the accrual of positive judgments to the agency about its technical capacity. Reputation management supposes an endogenous capacity to manage “competing views about technical expertise”, the “contestation as to what excellence in professional expertise implies” or the competition between agencies that claim scientific expertise in the same regulatory area (Christensen and Lodge 2018, p.122). These challenges are nearly inevitable in regulatory areas marked by uncertainty, politicization and ambiguity around scientific standards (Maor, 2007; McPhilemy and Moschella, 2019), as is broadly the case in the “regulatory science” that is used in areas of risk (Rushefsky, 1985; Jasanoff, 1990; Demortain, 2017), where the consideration of experiments, data or models are inseparable from the conflicting assumptions and values of the actors involved (Funtowicz and Ravetz, 1990).

The practical ways in which regulatory agencies handle the uncertainty surrounding what counts as expertise, and the ensuing challenges to their expert authority, are little explored in the literature on reputation. Technical reputation seems to be a capital that agencies trade for other forms of reputation, depending on the context – not something that is subject to competing conceptions. Expertise is generally approached as an apolitical, readily-interpretable or even measurable capability. Studies in the area consider how agencies utilize knowledge (Boswell, 2009; Weiss, 1979) as some sort of fixed commodity and resource that serves to send

unambiguous “professional signals” (Busuioc and Rimkute, 2020). In the available case studies of agencies’ choice towards science in risk controversies, the various scientific frameworks applied by regulatory agencies seem to operate like switches, which agency officials activate in reaction to external pressures (Lofstedt, 2011; Rimkute, 2018)¹. Controversy over what counts as “expertise” and political debates surrounding knowledge standards are rarely described in full and incorporated in the analysis. This may lead to problematic interpretations of agency actions, as either following a technical reputation (defined by reference to a *sui generis* type of science) *or* responding to alternative reputational concerns around its morality or performance². The fine-grain analysis of how and why an agency maintains a good technical reputation on an issue and not another (a frequent case) is limited as a result.

To fully understand the active construction of a technical reputation, we argue that a perspective considering how standards of authoritative knowledge are formed in the network of relationships between the agency and its audiences, and the action of agencies to assemble and lend credibility to these standards, represents a useful complement to the literature. To make progress in this direction, we associate Carpenter’s (2010) understanding of agencies’ reputation — as beliefs and symbolic constructs that emerge in dynamic networks of audiences — with notions borrowed from the field of science and technology studies that link the production of credibility to the formation of sociotechnical networks.

¹ Using a comparative perspective, Lofstedt (2011) for instance argues that risk regulation agencies in Europe will be more prone to use a hazard-based approach when no national firms are concerned with the product being assessed (in other words, they side with NGOs and apply a precautionary approach); whereas they will be more risk-based when national economic interests are at stake (i.e. side with industry and suggest that the risk is acceptable or non-existent).

² Rimkute’s (2018) argument that in the Bisphenol A controversy in Europe, the French regulatory agency (ANSES) was inspired by a reputation for public health protection, while the European Food Safety Authority by a technical reputation, encountered strong criticism from the informants themselves, as this interpretation neglects an important part of the work of ANSES towards science (see Genet, 2019).

Carpenter uses the notion of network of audiences to signify that agencies must work with multiple audiences that will understand, approve or contest its productions in varying fashions. They are constrained to demonstrate their expertness, by reference to the scientific standards that prevail in such network. Agencies may sometimes attempt to assemble a set of audiences to make sure their preferred expertise is foregrounded. As Moffitt shows, the strategic selection and enrolling of audiences through public consultation serves that purpose: “*Agency-structured public committees provide agencies with venues for deliberately revealing information to shape outsiders’ knowledge and perceptions, to define the boundaries of appropriate action (Jasanoff 1990), and to present a favorable agency image (Hilgartner 2000).*” (Moffitt, 2010, p. 883). This aligns with the use of the notion of network and audiences in science and technology studies (STS). In his work on stage management, Hilgartner emphasizes how members of a knowledge organization actively work to enroll audiences into a sociotechnical network that will judge positively the claims made by the organization (Hilgartner, 2000). He builds on seminal work in STS, to show that the success of a certain knowledge form rests first and foremost on the scientists’ capacity to network with allies and form its public (Shapin and Schaffer, 1985, p. 342). Scientific credibility depends on the social network in which the person, who is supposed to show that he or she knows, is embedded; a network that constitutes the basis of what may be called, following Shapin (1995), an economy of credibility.

We combine bureaucratic reputation research and science and technology studies to show that the technical reputation of an agency is indexed on the credibility that it attains in the network of audiences involved in judging its knowledge claims and expertise. Our framework is based on three supporting arguments. First, a network of audiences needs to be understood as operating on the basis of interdependent relations between different organizations. We therefore choose to speak of configurations of audiences – based on Elias’ concept of configuration

(1983) – to stress these interdependencies, and the overall equilibrium that may be reached when part or all of the components of the configuration are aligned. Second, audiences may align cognitively as it were, on a form of knowledge that will be judged convincing by them, to form a knowledge standard, or evidence. Third, agencies perform evidential work: they adjust or choose between knowledge standards, to align the audiences that dominate the construction of its reputation on a given issue. The technical reputation of an agency depends on the success of this cultural-cognitive management of audiences, and its ability to cultivate relationships with audiences (Carpenter, 2010; Moffitt, 2010; Maor, 2016) through evidential work, more than it results mechanically from the political economy of an issue or structural dominance of a given interest group.

In this context, reputation is the result of a dynamic process, a form of equilibrium around a particular standard of knowledge, which serves to uphold the agency's position in the complex and dynamic set of interdependencies in which it is engaged. A technical reputation is a judgment that emerges from a particular configuration of actors within which knowledge is produced and evaluated; a configuration, thus, within which agencies gain credibility by resorting to particular knowledge standards, or by altering these to produce knowledge that will be considered convincing – evidence – in this given configuration.

We illustrate this through an analysis of the shifts in EFSA's assessment of BPA. In this case, we see a correspondence between the rise of a public and controversy-prone set of organizations (more diverse groups, environmental in particular, with national perspectives), which EFSA handled through a public consultation, and an inclusion of broader considerations about the hazards of the substance, as well as a more realistic – that is, based on measurements, rather than simulations – exposure-driven calculation of the risks. This paper is based on information collected over many years of studying the activities of EFSA and ANSES (the French health

safety agency), through interviews, observations, participation in expert committees and archival research; as well as specific information about the case of BPA gathered in 2014 and 2015 through semi-structured interviews with scientific experts and agency officials, document analysis and attendance in the 2014 public meeting during which the new risk assessment of BPA was disclosed.

The risks of food contact materials: understanding shifts in expertise

The routine regulatory assessment of food contact materials

By law, the toxicity of chemical substances used in food packaging must be assessed, because they may migrate to the food itself and be ingested. EFSA is consulted to verify that the newly used substances do not transfer to food in quantities that endanger human health. In cases where migration may be important, EFSA uses a large set of studies (on absorption, distribution, metabolism and excretion; on reproduction in one species, and developmental toxicity in two species; studies on long-term toxicity/carcinogenicity, normally in two species, on top of *in vitro* tests and a 90-day study in rats (Silano et al. 2008)), performed under “Good laboratory practices” certification to ensure quality and replicability. The resulting data are used to compute synthetic figures about the substance. The first and main kind of computation is the Tolerable Daily Intake (TDI): the amount of the substance (in mg) a person can safely ingest each day during her lifetime. The TDI derives from the identification of the threshold below which the substance does not create any adverse effect in animals (the so-called NOAEL, for No Observed Adverse Effect Level), divided by 100 to cover the uncertainty surrounding the human’s potentially greater sensibility to the substance.

The methodology has evolved over the years to address the difficulty to identify a no-effect level in experimental data — a practice that involves a great amount of expert interpretation,

for instance concerning the appropriate methods of statistical extrapolation. Since experts can produce different interpretations, diverging judgements about the safety of the same substance from one agency to another, or sometimes inside one agency (Author, 2020), may emerge. Regulators have tried to minimize this inconsistency (EFSA 2009), because of the lack of objectivity it implies (Porter, 1992). A solution was found with the so-called Benchmark Dose (BMD) method, designed to produce a “reference point” or “point of departure”: the dose above which one can safely consider that toxicity will increase dramatically. The computation of the BMD is supported by several software programs. Its advantage is to force the evaluator to eliminate data sets that are unreliable, or tend to produce over-estimated NOAELs (Hardy et al., 2017). The TDI and BMD can be supplemented by a third conventional method, the Human Equivalent Dose (HED). To compute the latter, the BMD is multiplied by a factor that accounts for measurable differences between the body-surface-area of animals and humans, in replacement of the much criticized, conventional safety factors of 10, 100 or 1000. These computations are done following a meta-convention, known as the risk assessment-risk management framework (Author 2020). The dose calculations are combined with quantified information about the exposure of human populations to this hazard (i.e. an estimation of the doses that are found in “real life”), to produce a risk figure.

The standards of the TDI, BMD, HED and the risk assessment-risk management assemblage, all have a long, transnational history of formation. The TDI methodology traces back to the action of a number of toxicologists and medical doctors, in dedicated transnational symposia and the World Health Organization (WHO)(Author, 2011; Jas, 2015). The BMD was first designed for the purpose of the Environmental Protection Agency in the US in the 1980s (Crump, 1984), and then diffused through transnational networks of cooperation between regulatory scientists from both sides of the Atlantic. The HED, good laboratory practices

(GLPs) and the risk assessment framework have a comparable history of formation in transnational networks and international organizations – the Organization for Economic Cooperation and Development, the WHO and Codex Alimentarius (Author, 2012, 2014; Winickoff and Mondou, 2017).

Bisphenol A: the public controversy

The evaluation of food contact materials is thus informed by an ensemble of rules that were shaped transnationally and that provide scientific advisers and staff of regulatory agencies with legitimized and seemingly objective – because regular and replicable – ways of computing risks. Food contact materials, however, do not just exist as object of computation, but also as public risk object (Hilgartner, 1992). This is the case at least for certain, highly visible, substances like BPA.

An organic synthetic compound used in making plastics with poor water solubility, BPA is found in food contact materials, particularly beverage cans and bottles. From the early 2000s to 2015, the TDI for the substance hardly changed: it was set at 10 µg BPA/kg bw/day, based on a NOAEL of 5 mg/kg bw/day, identified in rat studies, and the application of a safety factor of 500 to compensate for the fact that there is no certainty that the substance is not toxic at low doses. This assessment was broadly shared by the FDA and EFSA (EFSA 2006, 2008). While there are many experimental studies of BPA available, this risk assessment long hinged on a limited set of studies (Tyl et al. 2002, Tyl et al. 2008) that were retained because of “*the large number of animals used [in these studies], the wide range of doses investigated and the fact that they were conducted according to GLP*” (Beronius et al., 2010, p. 138). The TDI translated into a statement that the substance posed no risk to the general population based on current knowledge.

EFSA's estimation of the TDI has been severely criticized by researchers studying endocrine disruption and by environmental groups. These actors have pressured the agency to recognize the low-dose effects of BPA on certain kinds of toxicity, particularly developmental toxicity. For EFSA, the studies demonstrating developmental toxicity at low dose of BPA were not replicated, and could not be fully considered valid. What is more, they were not conducted under GLP certification. For these reasons, agencies left these studies aside in their assessment (Beronius *et al.*, 2010, p. 144). For endocrinologists, this is particularly problematic and illegitimate, as it equates with ignoring potential problems of developmental toxicity, that only appear in low-dose, non GLP certified research that are, by all other scientific standards, of good quality (Myers *et al.*, 2008).

The risk controversy took on a political dimension when the substance came on the agenda of various national governments and NGOs in Europe. Members of the parliament or governments picked up on research establishing low dose risks of BPA, following already intense campaigning by environmental NGOs. This changing context reflected in the conduct of the regulatory risk assessment of BPA by EFSA.

In France, in 2008 and 2009, Bisphenol A became a common cause of concern, public debate and governmental activity. The decision by Canada to ban BPA in baby bottles was heavily reported in France. The scandal of the deaths attributed to the medicine Mediator, and accusations that France's medicine agency had overlooked signals and taken too long to withdraw the drug from the market, rendered the context sensitive. The BPA issue was the object of public action by a converging set of actors, from the NGO *Réseau Environnement Santé*, the socialist member of parliament (MP) Gérard Bapt, the Green party, highly successful in the 2009 European elections, and former Presidential candidate Ségolène Royale.

In this context, the food and environment agency (ANSES) was asked to update its assessment. Between 2008 and 2012, it produced five separate opinions, a full research note and two large expert reports. The French national health research institute (INSERM) was also asked to produce a systematic review of the effects of BPA on human reproduction. Environment health activists closely monitored each of these scientific efforts, commented on them publicly and mobilized MPs. In July 2009, the *Assemblée Nationale* put forward a proposal to ban BPA in plastic materials in contacts with food. The political context encouraged risk assessors to pay greater attention to even low risks associated with exposure to BPA and the protection of vulnerable populations. This also implied that BPA was addressed as a public concern, not in the regulatory framework of substance authorization, following pre-defined criteria. ANSES, in other words, was less constrained by international rules than EFSA, and more readily looked beyond the boundaries of strictly defined, standard regulatory knowledge. It was looking at BPA less as one substance to regulate among many, than as a unique substance specifically associated with marking, if not shocking, potential health effects on fetuses and newborn babies. In 2013, it declared that exposures to BPA during pregnancy posed a risk for the developing fetus (ANSES, 2013), taking into account studies that, whilst not having been published in peer-reviewed journals, produced evidence of possible health effects at low dose. ANSES broke the boundary of formally validated, GLP-certified and peer-reviewed studies, in a context where these were coming under increasing criticism for their limits, potential conflicts of interests and limited reach.

In Denmark, a center-right government placed endocrine disruptors on its agenda around the years 2007-2008, under pressure from a coalition of left parties in the Parliament that targeted BPA and pushed for a precaution-oriented regulatory effort at the European level. Several proposals for parliamentary resolutions and debates on the ban of BPA took place between 2009

and 2015. Despite disagreements among parties as well as between these and the government (some recognizing low-dose effects, others noting only that there remained uncertainties, which may be addressed through voluntary agreements with the industry), the presence of BPA in baby bottles and BPA's potential effects on children became a common concern. The Danish Technical University (DTU), operating under contract with the relevant ministries, re-assessed a number of studies in 2010. It found weaknesses in the GLP studies investigating reproductive toxicity effects, but noted that the hazards associated with BPA remained uncertain. In a political configuration centered on averting hazards and applying precautionary approaches, this was a sufficient basis to justify a national ban on BPA in baby bottles. In subsequent years, specifically during the period when social-democrats governed Denmark (2014-2015), there was a consensus for following the TDI calculated by the DTU – at $0.7\mu\text{g}/\text{kg}/\text{day}$ – rather than that of EFSA. The proponents of a new parliamentary resolution to phase out BPA discussed alleged conflicts of interest among EFSA's experts and the apparent preference of the agency to protect the industry rather than public health.

The French and Danish governments, with their respective risk agencies, were thus very active advancing a hazard-based assessment of BPA, in coherence with their willingness to restrict the use of BPA. These governments were not alone in Europe. The Danish Technical Agency had essentially expanded a preliminary assessment performed by a group of experts put together by the environmental agencies of the Nordic countries, notably the very active KEMI in Sweden, and of the Norwegian Scientific Committee for Food Safety (VKM), to attend to neurodevelopmental effects, and low dose effects more generally. Belgium had banned BPA in baby bottles too, following a scientific committee that also showed sensibility to arguments about the hazards of BPA, irrespective of the actual levels of exposure. Austria, in the same period, had reduced the presence of BPA in pacifiers and teething rings. The center-left Dutch

government of that time had not pronounced any restriction to the use of BPA, but several ministers had jointly mandated the scientific risk assessment institute, RIVM, to prepare an overview of the state of knowledge about BPA. These national moves were based on the same scientific argument: BPA is hazardous, particularly for developing organisms. And while exposure seems limited, or remains uncertain, the precautionary principle warranted European restrictions to the use of BPA.

EFSA and BPA: tinkering with conventions of risk assessment

These mixed science and policy positions against BPA were reflected at the European level. The environment committee of the European Parliament, notably, foregrounded BPA in a resolution of 2013 on endocrine disruptors. This resolution, propelled by the large group of Green MEPs resulting from the 2009 European elections and hailed by environmental NGOs, argued that European decisions on endocrine disruptors should consider the fact that “any exposure may entail a risk”, and use a hazard-based approach. The European Chemicals Agency had given suit to the French proposal to classify BPA as a presumed human reproductive toxicant. EFSA was requested by the European Commission to review the reports published by national agencies, notably the French and Danish ones in 2011.

Up to that moment, however, EFSA stuck to its initial assessment, against public accusations of neglecting the hazards, and the multiple alternative assessments in Member-States (see *Daily Express* 2007, *Pesticide & Toxic Chemical News*, 2007, *Food Chemical News* 2007, 2008). In 2012, EFSA committed to a re-evaluation of the toxicity of BPA and exposure to it. Simultaneously, EFSA’s management recognized the controversial nature of the issue, and the level of publicity and attention it got from industry groups and associations. Instead of treating

this as a routine risk assessment issue to be dealt with between the walls of an expert committee, it put to use a more active, public engagement policy. Since its creation, EFSA had developed a policy of engaging with “stakeholders”, to fully comply with its mission to communicate risks to the public, according to its founding regulation. Early on, it undertook to name representative organizations, from health and consumer NGOs to industry associations, with which it could engage on a more routine basis. It developed a practice of organizing workshops with stakeholders to discuss its risk assessments. The management of EFSA decided to use such a tool on the BPA issue, given the level of attention and publicity that now surrounded it. A public consultation was planned, partly online, the rest during a physical meeting. During that consultation, several key shifts in the assessment of risk were announced.

In the draft opinion it publicly released in 2014, the agency suggested to lower the TDI for BPA to 4µg/kg bw/d, without changing its overall assessment of the risk, concluding that “*there is no health concern for any age group from dietary exposure or from aggregated exposure*” (EFSA, 2014³). This risk assessment hinged on a number of adjustments to the set of conventional, transnational methods used for risk assessment. First, EFSA excluded a number of endpoints from the calculation of the TDI, particularly those observable at low-dose. This choice was based on a selected set of studies, constitutive of what the agency’s experts innovatively called a “line of evidence”, an approach that has similarities with the more internationally current “weight of evidence” approach. A line of evidence is constituted of several converging studies that “support the likelihood of toxicity at the endpoint”. The experts recognized that the effects on the mammary gland, kidney and liver were likely, but that reproductive and developmental toxicity were “as likely as not”. The opinion of the group of

³ A press release concomitantly argued that “there is no health concern” concerning BPA, with no further qualification. This announcement was later corrected.

experts mentioned and quantified those effects, but only as an element of “hazard identification”. In other words, the effects were not included in the computation of the TDI and of the final risk estimation. The effects on the mammary gland were simply considered through the inclusion of an additional uncertainty factor.

The lowering of the TDI resulted, second, from the choice to apply the HED methodology. The application of this “accepted” (in the words of a member of the expert group), but thus far unused, method was justified by the need to take into account the differences in the toxicokinetics of the substance in animals and humans. No default factor was needed, because the calculations were based on “real data” and measurements, not model-based inferences.

Third, EFSA considered that while the toxicity was perhaps higher than expected (hence the lower TDI), the estimated dietary exposure was “far lower” than what it had estimated in its 2006 opinion. The group of EFSA experts admitted a higher hazard, but decreased exposure, thus preserving their previous conclusion that the risk was, overall, absent.

The different moves decided by EFSA generated much debate during the public meeting held in March 2014 in Brussels. The representative of the DTU applauded EFSA for changing its TDI (though in part, sarcastically), but others asked why not lower it even more, and also questioned the very fact of applying the TDI methodology on a substance that does not display a clear toxicity threshold. Along with a number of other participants in the consultation, the DTU representative also questioned the way in which EFSA’s expert group formed a judgment of reliability on studies performed at low dose, and the reasoning behind the exclusion of many of these studies from the data considered in the final computation of the risk. The opinion, many thought, was not sufficiently explicit about the uncertainties that remained, and the possibility that the expert group could be wrong.

The members of the expert group of EFSA, along with the chair of the Scientific Committee of the agency, reiterated the legitimacy of what they considered an inclusive approach to regulatory evidence. On the one hand, they applied what they described as a “filtering mechanism”, selecting only the most “reliable” or “stronger” studies to compute the TDI. On the other hand, other studies were not left out entirely. They were used to characterize the hazard. This did not please entirely the representatives of the DTU and of the NGO CHEM Trust, who argued that the exclusion of certain studies, hence of certain endpoints like developmental toxicity from the TDI calculation, was hard to accept. In response, the experts conceded that the final opinion will map even more explicitly these “uncertainties”. The line was fine between giving in to pressure to consider chronic, low-dose effects, and staying faithful to the methods that assure its objectivity but reveal limited hazards, as the exchanges during the public meeting showed. But the opinion that was eventually published in January 2016 following this public meeting, closed the controversy for a moment, as far as the European stage is concerned at least⁴.

The organization of evidence

The changes brought by EFSA to its BPA assessment in 2014 constitute a particularly interesting episode to interpret. The 2014 opinion was forged by a number of preliminary methodological choices, and some tinkering with transnational conventions for risk assessment. This in turn led to the organization of a working group to establish a new framework. Among the key changes, the use of the HED methodology; the notion of “line of evidence”; and an

⁴ Since then, the EFSA established an international working group to create a detailed and transparent framework to guide the assessment of the substance, continuing the ad hoc development of methodology to review BPA studies. It committed to re-evaluate BPA in 2018, to better examine studies released in the meantime.

adjusted classification of uncertainties, which justified applying an *ad hoc* “filtering” of studies, at the point of passage between the hazard identification and hazard characterization stages of the assessment. Finally, EFSA used to the fullest possible extent the notion that a risk assessment balances a hazard characterization and an exposure assessment, lowering its exposure estimate to compensate for the slightly increased estimation of the hazards.

These adjustments show that EFSA did not simply reject a risk-based approach, combining hazard and exposure information, to embrace a pure, hazard-based approach that NGOs and a handful of governments acting against BPA in baby bottles favoured. It incorporated, in the suite of data and studies considered for the risk assessment, a larger set of knowledge to represent the views of emerging audiences, in scientific work. It both gestured towards the kind of knowledge that this set of audiences judged more credible, but stuck to the overall method of risk calculation (including data on exposure), to preserve the consistency of a risk assessment according to which BPA does not represent a major threat to public health. Preserving its reputation as a technical agency comes down to performing this subtle work on the science – convincing a particular set of audiences by considering the data and studies that they favour, while maintaining a level of consistency with past assessments, through an *ad hoc* demonstration of the absence of the risk. This evidential work neatly reflects the changing configuration of audiences that EFSA engages with.

Indeed, agencies such as EFSA face a structural tension, due to the variety of missions which they pursue, and to the corresponding audiences with which they engage. EFSA is in relation, first, with industrial audiences, since it contributes to the regulation of an industrial sector and its productions through the process of risk/product assessment. An important motivation in EFSA’s creation was the need to establish an agency in capacity to perform the work of evaluating the hazards and risks of regulated products, the diversity of which increased since

2000, with the adoption of more laws requiring product assessment. Second, EFSA is embedded in a network of other agencies with which it is in constant relation and that engage with its assessments. Indeed, EFSA has a cross-European coordinating mission and duty to enhance the harmonization of scientific risk opinions (Vos, 2000). Third, EFSA is also in close interaction with environmental and health activists, as well as consumer associations. Ever since its creation, it has established close connections with these organizations, notably the most representative ones, through a formal body and the organisation of frequent public consultation events.

These audiences may apply different criteria of evidence, and trust different data elements. It is obviously true of the industry and of member-states, who directly contribute scientific knowledge to EFSA. But it is also true of NGOs and other civic groups. While initially more active at the national level, NGOs now engage with EFSA regularly, criticizing its scientific assessments, down to its selection of studies or ways of interpreting data and computing hazards and their frequencies. This tension among audiences reflects in the scientific and epistemic operations of the agency and its experts.

The history of the methodologies that were applied for the assessment of the hazards of BPA shows that risk assessment mainly followed transnational conventions, that took shape among representatives of the first two audiences and were embedded in distinct organizational forms. Tracing the emergence of these methodological and regulatory concepts reveals the existence of a transnational social formation, which may be described as an invisible college (Author, 2011), comprising agency scientists, industrial scientists and academics interested in risk assessment. The emergence of regulatory methodologies in such a transnational network is driven by standardization and the application of internationally-agreed-upon rules (Cambrosio et al., 2006). The transnational methods of regulatory assessment of products – in this case TDI,

BMD, HED – are forged under the influence of the notion that regulatory agencies of different countries cannot possibly follow different methods, because inconsistency in scientific assessments among regulatory agencies invites criticism and controversy, and alters legitimacy.

What the case of BPA reveals, however, is that the configuration dramatically changed in the first part of the 2010s. A new alignment of a handful of member-states, national or European environmental NGOs, as well as politicians in the European Parliament, resulted in alternative sets of information, data and studies gaining greater weight. With such audiences, the dominant epistemology grants less value to compliance with standard methodologies and to a mechanical kind of objectivity. These audiences emphasize the collective review of facts and studies, joint selection of data, as well as *in situ* exploration of the resulting computations of the risk and remaining uncertainties. In this more deliberative mode, the interpretation of regulatory knowledge is performed competitively, on public stages with their own rules and frames of action, under the influence of various, often explicit interests and worldviews (Hilgartner, 2000; Callon et al., 2009).

We see the case of food contact materials and BPA as representative of a broader tension that underpins the work of regulatory agencies dealing with health and environmental risks, reflected here in EFSA's evidential work. The knowledge that was articulated in the course of the assessment of BPA reflects the engagement with various sets of audiences over time, and their epistemological expectations. The latest assessment of BPA risks by EFSA, and the shifts that the agency and its key experts conceded, demonstrate the fact that, initially, it has taken great care in mastering and applying transnationally-agreed methods, inscribed in guidelines, of particular importance in the relationship with regulated entities and other national agencies. These risk-based methodologies, which combine information about the hazard and probabilistic judgements based on exposure data, were a key component of the

credibility of the agency. The fact that they had been established in organizations removed from public scrutiny made the engagements of the different actors *vis-à-vis* each other and towards commonly agreed standards binding. Their adaptation, adjustment or outright reinvention in the process of assessing the substance, reflects the growing weight of other audiences, and the need of EFSA to engage with them and their demand for more investigation and deliberation around the existence of hazards, separately from exposure data. With the shift enacted in 2014, which the constitution of a network of audiences and their joint presence at the 2014 public consultation meeting enabled, EFSA moved from an algorithmic mode of assessment – one in which data considered sequentially informs the application of a criterion – towards a more demonstrative mode, publicly staging the way in which it formed key judgments. The shift from a notion of “weight of evidence” to “line of evidence” best illustrates this. Rather than mechanically weighing each pieces of evidence, EFSA’s experts actively constructed the evidence base, making sure that the data and computations put forward by NGOs and Member-states that judged BPA hazardous, were represented in this knowledge base; however not to the point of entirely reversing its overall assessment of the risk associated to BPA.

A positive technical reputation as an ‘expert body’, will result from the successful construction of evidence with the actual audiences of an organization. Approaching this evidential work reveals an important tension in the work of regulatory agencies. Upstream cultural engagement with transnational audiences of regulated industries and other regulators, resulting in the fact of embracing a form of *a priori* regulatory science, can come in contradiction with the agencies’ engagement with other, more situated audiences that respond to an emerging controversy. This contradiction is at the heart of regulatory knowledge. European regulatory agencies work at the intersection of several missions (delivering

regulatory opinions in view of the commercialization of products, coordinating national agencies or harmonizing their assessments, disclosing information to the public), address distinct audiences (the transnational regulated industry and expert communities, national agencies, NGOs), and are thus engaged in configurations that differ markedly from the point of view of the knowledge that is considered conclusive.

The case of BPA shows that agencies such as EFSA sometimes adapt the standards by which the existence of a hazard or risk is proven. A configuration of audiences may vary strongly from one regulatory issue or controversy to another, even within the same agency. The case of BPA described above illustrates a situation in which the interactions with a transnational audience of experts specialized in the forging of regulatory assessment methodologies, were progressively balanced out by interactions with European NGOs and national risk assessors, resulting in the local adaptation of transnational conventions.

Conclusion

This paper has focused on the work regulatory agencies perform to convince audiences of the existence or absence of a risk, along with the credibility of their calculations. Against the tendency to reduce this work to the mechanical application of methods and a predetermined knowledge style on which the agency's technical reputation rests, we emphasize how this reputation is the result of evidential work: i.e. a practice of forging scientific demonstrations in the configuration of audiences that dominate an issue. The paper shows that the more a risk issue is politicized, and the more audiences advance particular knowledge claims regarding these risks, the more an agency is likely to engage in the adaptation of its standards of regulatory evidence to match these claims. The knowledge claims that appear credible, then, are those that are produced with a set of audiences, in such a way that a positive scientific reputation will be associated with a more deliberative, rather than technocratic, style of

engagement. Accordingly, the regulatory knowledge that an agency like EFSA produces is the outcome of the work carried out to manage the tensions within configurations of audiences, and their respective visions of what constitutes the right mode of regulatory knowledge for health risks.

We have used the enduring controversy over the risks of BPA to illustrate this evidential work. This episode points to the existence of a phenomenon – the interactional evidential work performed in configuration of audiences – that concurs with the notion that reputation is a complex construct, and that no audience has full control over reputation. The episode reveals a larger trend, observable over the last decades, in which multiple stakeholders engage with science on a regular basis, with profound consequences on regulatory agencies' capacity to operate and maintain their technical reputation, and the associated images of scientific excellence, objectivity or neutrality. The conditions in which regulatory agencies with a scientific mission gain legitimacy has evolved over the years, as science has become an object and resource of political work for a greater number of audiences than was the case when these agencies were first created. The capacity of NGOs to engage with risk assessments and their technical requirements has increased, as has their ability to deconstruct the methods used and to suggest alternative approaches. Agencies initially reacted to these controversies by opening up to activist groups and allowing them to voice their concerns and push for alternative modes of risk assessment. All regulatory agencies that rely on technical expertise for their mission have experienced this shift. Drug agencies were the first to understand and adapt to this shift, notably in relation to Aids and later rare diseases. Agencies in charge of assessing food products, chemicals or workplace safety have also learned to work with extended networks of audiences.

The practice of regulatory science requires working with sets of audiences – experts, NGOs, the regulated industry, and so on — and acknowledging the consequences of these multiple interactions on the definition of credible knowledge. This situation decisively constrains the capacity of agencies to appear as experts and designate a single body of knowledge as a reference point. The ways of managing the tensions inherent in these complex science-mediated relationships with audiences vary across agencies. EFSA, for instance, remains a steadfast defender of transnational conventions, and displays a limited willingness to alter these norms, at least for certain issues it is responsible for. One reason may have to do with the fact that experts who sit on EFSA committees have taken part in the elaboration of these conventions. Another relates to the position of the agency *vis-à-vis* its national counterparts and the need to ensure that all these organizations follow the same procedures. In France, ANSES questions more frequently standard protocols, possibly to appear tough on industry, but also given the presence at the head of the agency of several directors with an environmental or occupational health background. Yet other cases may be observed. Working with extended, and often contentious, networks of audiences may lead to a strategy of internal segmentation, with each agency devoting specific and compartmentalized resources to working with distinct audiences. But it could also converge around a more pluralistic mode of decision-making, in which the various points of view and arguments are presented and discussed. This in turn suggests that the boundary between science and politics has become a topic of organizational reform, either within or between agencies, as these learn to work with and navigate within increasingly complex networks of audiences, while still upholding their reputation and apparent neutrality.

Our case substantiates the claim that reputation varies from one issue to another, as do configurations of audiences. It confirms the results of work on regulatory capture, particularly

the fact that agencies are rarely fully aligned with either public or private interests (Carpenter and Moss, 2014), and actively “balance” audiences (Ligboer, 2020). It is the knowledge order that results from evidential work and from what agencies do, which may support a qualification of the agency as aligned with one or another interest, or a coalition of interests (Dedieu, forthcoming). The originality of our paper is to show that agencies working with science perform the same sort of strategic, interactional and audience-driven work that regulatory agencies perform more generally, be it through communication or the law. Regulatory agencies are neither fully in control over regulatory science — as is the FDA, an agency with a large “conceptual power”, or power over the norms of regulatory science (Carpenter, 2010) — or condemned to embrace knowledge advanced by some of their audiences. In between these two extremes, we have revealed the amount and richness of the work that agencies perform on and through science, to forge their power. Given the variety of agencies that need to collect technical information from multiple audiences as part of their mission, evidential work is certainly an integral part of the repertoire of strategic, legitimacy-enhancing actions of regulatory agencies in general.

References

Author 2011, 2012, 2014.

Beronius, A., Rudén, C., Hakansson, H. & Hanberg, A. (2010). Risk to All or None?: A Comparative Analysis of Controversies in the Health Risk Assessment of Bisphenol A. *Reproductive Toxicology*, 29, 2, 132–146.

Busuioc E.M. (2016). Friend or Foe? Inter-Agency Cooperation, Organizational Reputation, and Turf. *Public Administration*, 94, 1, 40-56.

Busuioc M. & Rimkutė D. (2020a). The promise of bureaucratic reputation approaches for the EU regulatory state. *Journal of European Public Policy*, 27, 8, 1256-1269. <https://doi.org/10.1080/13501763.2019.1679227>

Busuioc M. & Rimkutė D. (2020b). Meeting expectations in the EU regulatory state? Regulatory communications amid conflicting institutional demands. *Journal of European Public Policy*, 27, 4, 547-568. <https://doi.org/10.1080/13501763.2019.1603248>

Callon, M, Lascoumes, P. & Barthe, Y. (2009). *Acting in an Uncertain World: An Essay on Technical Democracy*. Cambridge, MA: MIT Press.

Cambrosio, A., Keating, P., Schlich, T. & Weisz, G. (2006). Regulatory Objectivity and the Generation and Management of Evidence in Medicine. *Social Science & Medicine* 63, 1, 189–99.

Carpenter D.P. & Krause G.A. (2012). Reputation and Public Administration, *Public Administration Review*, 72, 1, 26-32.

Carpenter, D. P. (2001). *The Forging of Bureaucratic Autonomy : Reputations, Networks, and Policy Innovation in Executive Agencies, 1862-1928*. Princeton University Press.

Carpenter, D. P. (2010). *Reputation and Power: Organizational Image and Pharmaceutical Regulation at the FDA*. Princeton, NJ: Princeton University Press.

Carpenter D.P. & Krause G.A. (2012). Reputation and Public Administration, *Public Administration Review*, 72, 1, 26-32.

Carpenter, D. and Moss, D. (Eds)(2014). *Preventing regulatory capture: special interest influence and how to limit it*. Cambridge University Press.

Christensen, T. & Lodge, M. (2018). Reputation Management in Societal Security: A Comparative Study. *The American Review of Public Administration*, 48, 2, 119-132.

Crump, K.S. (1984). A New Method for Determining Allowable Daily Intakes. *Toxicological Sciences* 4, 5, 854–71.

Daily Express (2007). *Revealed: The killer chemicals in our food packaging*. 30 January 2007.

Dedieu, F. (2021). Organized Denial at Work. The Difficult Search for Consistencies in French Pesticide Regulation. *Regulation & Governance*, 10.1111/rego.12381, Early Online View.

Demortain, D. (2017). Expertise, Regulatory Science and the Evaluation of Technology and Risk: Introduction to the Special Issue. *Minerva*, 55, 3, 139-159.

EFSA (2006). *Opinion of the Scientific Panel on Food Additives, Flavourings, Processing Aids and Materials in Contact with Food on a request from the Commission related to 2,2-bis(4-hydroxyphenyl)propane (bisphenol A)*. European Food Safety Authority, Parma.

EFSA (2008). Toxicokinetics of Bisphenol A. Scientific Opinion of the Panel on Food additives, Flavourings, Processing aids and Materials in Contact with Food (AFC). European Food Safety Authority, Parma.

EFSA (2009). Use of the benchmark dose approach in risk assessment: guidance of the Scientific Committee. *The EFSA Journal* 1150, 1-72

EFSA (2014). Draft Scientific Opinion on the risks to public health related to the presence of bisphenol A (BPA) in foodstuffs. *The EFSA Journal*, un-numbered, available at <https://www.efsa.europa.eu/sites/default/files/consultation/140117.pdf>

Elias, N. (1983). *The court society*, Oxford, Blackwell.

Food Chemical News (2007). EFSA panel "relies quite heavily" on industry data. 9 April 2007.

Food Chemical News (2008). EFSA dismisses Health Canada concerns over Bisphenol A. 28 July 2008.

Funtowicz, S. O., & Ravetz, J. R. (1993). Science for the post-normal age. *Futures*, 25, 7, 739-755.

Genet, R. (2019). Regarding the publication of Dovilė Rimkute, 'Organizational reputation and risk regulation: The effect of reputational threats on agency scientific outputs' (*Public Administration*, 96(1), 70–83). *Public Administration*, 97, 1, 235-236.

Ginsburg, R. (1997). Quantitative Risk Assessment and the Illusion of Safety. In *Work, Health, and Environment: Old Problems, New Solutions*, edited by C. Levenstein & J. Wooding. New York, London: The Guilford Press, 8-15.

Hardy, A., Benford, B., Halldorsson, T., Jeger, M.J., Helle Knutsen, K., More, S., Mortensen, A., & al. (2017). Update: Use of the Benchmark Dose Approach in Risk Assessment. *EFSA Journal* 15 1, e04658.

Hilgartner, S. (1992). The Social Construction of Risk Objects. In *Organizations, Uncertainties and Risk*, edited by J. Short Jr & L. Clarke. Boulder: Westview Press, 39–53.

Hilgartner, S. (2000). *Science on Stage: Expert Advice as Public Drama*. Stanford: Stanford University Press.

Jas, N. (2013). Adapting to 'reality': The Emergence of an International Expertise on Food Additives and Contaminants in the 1950s and Early 1960s. In *Toxicants, Health and Regulation since 1945*, edited by S. Boudia & N. Jas. Pickering & Chatto, 47–69.

Jasanoff, S. (1990). *The Fifth Branch: Science Advisers as Policymakers*. Boston, MA, London: Harvard University Press.

Libgober B. (2020). Strategic Proposals, Endogenous Comments, and Bias in Rulemaking. *The Journal of Politics* 82, 2, 642–656.

Lofstedt, R. E. (2011). Risk versus Hazard – How to Regulate in the 21st Century. *European Journal of Risk Regulation*, 2, 2, 149-168. <https://doi.org/10.1017/S1867299X00001033>

Maor, M. (2007). A Scientific Standard and an Agency's Legal Independence: Which of These Reputation Protection Mechanisms Is Less Susceptible to Political Moves?. *Public Administration*, 85, 4, 961-978.

Maor, M. (2011). Organizational reputations and the observability of public warnings in 10 pharmaceutical markets. *Governance*, 24, 3, 557-582.

Maor, M. (2015). Theorizing Bureaucratic Reputation. In A. Wæraas & M. Maor (Éds.), *Organizational Reputation in the Public Sector*. Routledge, 17-36.

Maor, M. (2016). Missing Areas in the Bureaucratic Reputation Framework. *Politics and Governance*, 4, 2, 80-90.

McPhilemy S., Moschella M. (2019). Central banks under stress: Reputation, accountability and regulatory coherence. *Public Administration*, 97, 3, 489-498.

Moffitt, S.L. (2010). Promoting Agency Reputation through Public Advice: Advisory Committee Use in the FDA. *The Journal of Politics*, 72, 3, 880-893.

Myers, J.P., vom Saal, F.S., Akingbemi, B.T., Arizono, K., Belcher, S., Colborn, T., Chahoud, I. & al. (2009). Why Public Health Agencies Cannot Depend on Good Laboratory Practices as a Criterion for Selecting Data: The Case of Bisphenol A. *Environmental Health Perspectives* 117, 3, 309-315.

Pesticide & Toxic Chemical News (2007). EFSA's assessment of Bisphenol A draws fire: research on Bisphenol A accused of not being authentic. 19 February 2007.

Porter, T. (1992), "Objectivity as standardization: The rhetoric of impersonality in measurement, statistics, and cost-benefit analysis". *Annals of Scholarship* 9, 19-59.

Rimkutė, D. (2018). Organizational reputation and risk regulation: The effect of reputational threats on agency scientific outputs. *Public Administration*, 96, 1, 78-83.

Rushefsky, M. E. (1986). *Making Cancer Policy*. Albany, New York: SUNY Press.

Séralini, G.-E., Clair, E., Mesnage, R., Gress, S., Defarge, N., Malatesta, M., Hennequin, D. & Spiroux de Vendômois, J. (2014). *Republished Study: Long-Term Toxicity of a Roundup*

Herbicide and a Roundup-Tolerant Genetically Modified Maize. *Environmental Sciences Europe* 26, 1, 14.

Shapin, S. & Schaffer, S. (1985). *Leviathan and the Air-Pump: Hobbes, Boyle and the Politics of Experiment*. Princeton, N.J: Princeton University Press.

Silano, V., Bolognesi, C., Castle, L., Cravedi, J.-P., Engel, K.-E., Fowler, P., Franz, R. et al. (2008). Note for guidance for the preparation of an application for the safety assessment of a substance to be used in plastic food contact materials. *The EFSA Journal* 6, 7, 1-41.

Winickoff DE, Mondou M (2017). The problem of epistemic jurisdiction in global governance: The case of sustainability standards for biofuels. *Social Studies of Science*, 47, 1, 7–32.