Regulation & Governance

Regulation & Governance (2021) 15, S123–S142 doi:10.1111/rego.12378

Why meta-research matters to regulation and governance scholarship: An illustrative evidence synthesis of responsive regulation research

Jeroen van der Heijden

School of Government, Victoria University of Wellington, Wellington, New Zealand; and School of Regulation and Global Governance (RegNet), Australian National University, Canberra, Australian Capital Territory, Australia

Abstract

This article is an introduction to meta-research, a systematic and replicable process of synthesizing research findings across a body of original research. After introducing the reader to the core of meta-research methodology, meta-research logic and tools are applied to present an evidence synthesis of empirical research on responsive regulation. The article concludes with a meta-research agenda for regulation and governance scholarship, and five key lessons from the empirical responsive regulation literature.

Keywords: evidence synthesis, meta-research, responsive regulation.

1. Introduction

"We are drowning in information, while starving for wisdom."

(Wilson 1999, p. 294)

The growth in the size and diversity of regulation and governance scholarship warrants it being subjected to meta-research questions. For example, to what extent can we aggregate existing knowledge from individual regulation and governance studies so that the whole of the knowledge becomes more than the sum of its parts? Or what explains the variance of insights presented in regulation and governance studies resulting from empirical studies on similar topics? The aim of meta-research (also known as metascience) is "to generate a more comprehensive and trustworthy picture of the topic being studied than is possible from individual studies" (Gough *et al.* 2012, p. xiii). Meta-research seeks to create new knowledge from separate, existing studies in a systematic, transparent, and replicable manner.

Since the early 2000s, there has been rapid growth in meta-research across the social sciences. Publications are now easily accessible through databases, and emerging technologies ease systematic analysis (Bearfield & Eller 2008; Schooler 2014; Sutton *et al.* 2016). However, more is at stake. The social sciences are facing a "credibility crisis" (Freese & Peterson 2018; Hardwicke *et al.* 2020). Social science theories and concepts are often "essentially contested" and allow for considerable freedom of application, and this (combined with often "noisy" social science research settings – i.e. treatment heterogeneity, measurement error, and contextual confounders) may result in problems such as flawed operationalizations and measurements, exaggerated estimates of effect sizes, and misinterpreted causality (Loken & Gelman 2017; National Academies of Sciences 2019). Publication biases, academic reward structures, and academics' lack of training in meta-research methods may further undermine the extent to which social science research can be

Correspondence:Jeroen van der Heijden, School of Government, Victoria University of Wellington, 23 Lambton Quay, Pipitea Campus, Wellington 6011, New Zealand. Email: jeroen.vanderheijden@vuw.ac.nz

* Conflict of interest: Jeroen van der Heijden receives funding from the New Zealand Government Regulatory Practice Initiative (G-REG) and kindly thanks them for their financial support. He trusts the G-REG support has not affected his objectivity in writing the current article.

[The copyright line for this article was changed on 07 October 2021 after original online publication]

Accepted for publication 16 December 2020.

aggregated and reproduced (Cooper 2017; Gurevitch *et al.* 2018). Scholars of regulation and governance often face similarly open theories, research environments, and publication cultures.

To present the relevance and basics of meta-research for regulation and governance scholarship, this article presents an overview of the state of the art in this discipline, building on an exploration of a set of meta-research handbooks published since 2009 and the related literature. The overview introduces the reader to the following: the different types of meta-research; the different foci of meta-research (e.g. the quality of research activity in a given area, the application of a theory, or evidence of an intervention); and the typical flow of carrying out meta-research and the best practices for its various stages (e.g. collection of source documents, setting inclusion/exclusion criteria, analyzing data, interpreting results, and presenting findings). The overview is by no means intended as an exhaustive discussion of the meta-research literature, and readers interested in pursuing their own meta-research projects are strongly advised to consult the various sources on which this article is built.

Informed by this overview, and by way of illustration, the article presents a realist evidence synthesis of published responsive regulation research (building on the work of Ayres and Braithwaite [1992]). The synthesis presents a map and a gap analysis of responsive regulation research, a critical appraisal and quality assessment of the publications included in the synthesis, and knowledge aggregation that moves beyond the insights from individual publications. The responsive regulation synthesis builds on a set of 65 publications (all empirical studies) that were selected in a systematic, transparent, and replicable manner from an initial database of 1,012 original peer-reviewed journal articles. The synthesis applies the meta-research methodology discussed in the first half of the article, and online appendices give insight into all the steps, from source document selection to analysis, including lists of the codes and keywords used.

The article concludes with a brief meta-research agenda for regulation and governance scholarship, a call on the editors of *Regulation & Governance* to embrace meta-research in future issues of the journal, and five key lessons from the empirical peer-reviewed responsive regulation literature.

2. What is meta-research?

The image sketched above of a credibility crisis is not unique to the social sciences. The natural sciences and humanities have long been subject to similar challenges (Wilson 1999). The challenges are particularly tricky for the social sciences and the humanities, however. In the natural sciences, replication of studies is often possible, but this is generally not the case in the social sciences (and the humanities) – regulation and governance scholarship included. In a nutshell, a replication study seeks to assess the robustness of originally published research findings. Typically, a different group of scholars follows the same protocols, methodologies, and research design as the original group of scholars, and subjects a similar object of study to a similar intervention. Positive replication validates the original findings. Failure to replicate indicates, at best, that more research is required to understand the phenomena of interest, and, at worst, that the original findings were fraudulent (Braver *et al.* 2014; Schooler 2014).

In the social sciences, such direct replication is often impossible to achieve (and it may be questioned whether this "golden standard" is desirable at all). Many phenomena cannot be replicated because too many factors and variables are at play (Even 2002; Block & Kuckerts 2018). How, for example, are the development of regulatory interventions, the performance of regulatory regimes in situations of crisis, or changes in the politics of regulation over time to be replicated? Also, direct replication comes with (other) technical challenges that can affect the study findings. A small, unnoticed change between the original and the repeat study setting might have a major impact on the findings of the repeat study. It is for this reason that scholars in the social sciences have begun to explore and apply another tool from the toolbox of natural sciences to address the credibility crisis: meta-research.

Meta-research should, however, not be considered a (mere) second-best option when replication studies are not possible. Meta-research is a true equivalent to replication studies (Booth *et al.* 2016; Heyvaert *et al.* 2017). Like replication research, meta-research is a form of "research-on-research" (Hardwicke *et al.* 2020). Meta-research is *a systematic and replicable process of synthesizing research findings across a body of original research* (see also Ringquist 2013; Eklund Karlson & Takahashi 2017). The roots of meta-research can be traced back to the 17th century, but it was not until the start of the twentieth century that it gained traction – first in the natural sciences (led by medicine) and much later in the social sciences and humanities (for the historical developments

Table 1 An overview of meta-research handbooks†

Handbook title	Orientation			
	Qual	Quant	Both	Mixed
The Handbook of Research Synthesis and Meta-Analysis (Cooper *et al.* 2019)			X	
Research Synthesis and Meta-Analysis: A Step-by-Step Approach (Cooper 2017)			X	
Using Mixed Methods Research Synthesis for Literature Reviews (Heyvaert *et al.* 2017)				X
A Resource for Developing an Evidence Synthesis Report for Policy-Making (Eklund Karlson & Takahashi 2017)			X	
Systematic Approaches to a Successful Literature Review (2nd ed.) (Booth *et al.* 2016)			X	
Applied Meta-Analysis for Social Science Research (Card 2015)		X		
Meta-Analysis for Public Management and Policy (Ringquist 2013)		X		
An Introduction to Systematic Reviews (Gough *et al.* 2012)			X	
Systematic Synthesis of Qualitative Research (Saini & Shlonsky 2012)	X			
An Introduction to Qualitative Research Synthesis: Managing the Information Explosion in Social Science Research (Howell Major & Savin-Baden 2010)	X			
Introduction to Meta-Analysis (Borenstein *et al.* 2009)		X		

†Qual = focus mainly on qualitative tools and data; Quant = focus mainly on quantitative tools and data; Both = focus on quantitative and qualitative tools and data; Mixed = focus on mixed method tools and data.

of meta-research, see, among others, Borenstein *et al.* [2009] and Shaddish and Lecy [2015]). It is only since 2000 that meta-research has begun to gain a foothold in areas such as public policy and public management (Ringquist 2013; Gurevitch *et al.* 2018).

Meta-research started off as a branch of science that used of statistical tools to synthesize findings (particularly average effect sizes) across quantitative studies (Borenstein *et al.* 2009; Card 2015). Since then, meta-research has developed as an inclusive form of research and provides a logic, methods, and guidelines for the use of quantitative, qualitative, and mixed-method tools to synthesize findings across qualitative, qualitative, and mixed-data studies (Gough *et al.* 2012; Heyvaert *et al.* 2017). Indeed, the breadth and depth of meta-research in the social sciences becomes clear by looking at, for example, publications in peer-reviewed journals such as *Systematic Reviews* and *Research Synthesis Methods* or the coverage of the meta-research handbooks that have appeared over the last decade – see Table 1 for a snapshot. More and more commonly, meta-research is applied to study solely qualitative source data, to supplement syntheses of quantitative source data with syntheses of qualitative source data, and to combine qualitative and quantitative data in mixed syntheses.

With the ongoing development of meta-research, its applicability has expanded. Besides the assessment of average effect sizes, scholars are now using it to answer questions such as: How, when, why, and under what conditions does a theoretical proposition hold (Cooper 2017)? How have theories, opinions, and values, or people's understanding of phenomena, developed (Gough *et al.* 2012)? What is the broad and comprehensive picture that emerges across a large number of study outcomes (dozens to hundreds) (Gurevitch *et al.* 2018)? What new knowledge can be created from separate existing elements, and what do the elements mean as a collective body of knowledge (Howell Major & Savin-Baden 2010)? What are the most promising avenues for future research (Cooper *et al.* 2019)? What explains the variance in empirical results across original studies (Ringquist 2013)? How do the disparate elements of the phenomena of interest connect and interact (Saini & Shlonsky 2012)? How biased is a body of knowledge, and what is the quality of a body of knowledge (Mayo-Wilson & Grant 2019)? What practical or usable lines of (policy) action have been provided within a body of knowledge (Heyvaert *et al.* 2017)?

To round up this very brief overview, meta-research is an active field of scholarship. It is subject to ongoing developments and debates about which approaches are best applied in which situations, and why, and when. There is a growing consensus that scholars should aim for a systematic, transparent, and replicable research synthesis when this is possible (Booth *et al.* 2016). Also, meta-research can be applied not "just" to theory testing, but also to mapping, exploring, and interrogating how policy debates have emerged and changed over time, or how scholars have engaged with analytical concepts across time and space (Cooper *et al.* 2019).

3. The basics of meta-research

At this point, the critical reader may ask: "But are those questions not the questions you would expect to see answered by a traditional literature review? What is so special about meta-research?" It is, after all, safe to assume that everyone reading this article has some experience of undertaking syntheses of existing knowledge bases, or at least has been exposed to their results – the journal *Regulation & Governance*, for example, has published a range of review and synthesis articles since its initial issues (Van der Heijden 2010; Coglianese 2011; Short 2021). However, not all knowledge syntheses are created equally. In what follows, the research handbooks highlighted in Table 1 will be used as the foundation of a synthesis of their main line of reasoning about how to carry out meta-research, and how such meta-research differs from "traditional" literature reviews.[1]

Traditionally, literature reviews in the social sciences have been carried out in a narrative manner (Cooper 2017). Such narrative reviews often lack explicit standards of proof, standards of evidence, and standards for synthesizing their findings (Melendez-Torres *et al.* 2017). They also do not (claim to) review all the literature available or provide criteria for the inclusion and exclusion of sources, they do not (claim to) check the reliability of source material, and they do not (claim to) follow a systematic and replicable approach to extract insights from the source material (Gough *et al.* 2012). In short, narrative reviews and meta-research serve different purposes. Narrative reviews are useful for exploring ideas and advancing conceptual frameworks, but they are not typically meant to provide an accurate summary and synthesis of findings across studies (Gurevitch *et al.* 2018).

3.1. Leaving a trail for others to follow

Meta-research for the social sciences is often conceptualized as following "one of two general approaches: combining and comparing studies" (Card 2015, p. 17). Such combining or comparing, or both, can be done in an aggregative, an integrative, or an interpretive manner (Saini & Shlonsky 2012; Heyvaert *et al.* 2017). Aggregative meta-research typically results in (quantitative) average effect sizes and the (qualitative, quantitative, or both) prevalence of themes across findings. Integrative meta-research typically results in (new) taxonomies, (new) concepts, and (new) descriptions of phenomena across studies. Interpretive meta-research, finally, typically results in new inductive understandings that move beyond the findings of individual studies and that may provide a better understanding of how different elements of the phenomena of interest connect and interact. However, meta-research is more about the *how* than about the *what* of such syntheses and comparisons – that is, it is about how to undertake them.

From an overview of the handbooks of meta-research identified in Table 1, as well as the broader literature on the topic, it becomes clear that meta-research follows a stepwise approach. There is no set number of steps to follow. Differences in ontological and epistemological assumptions, the reasons for undertaking the meta-research, and the strategies for source collection and analysis have resulted in different approaches to producing high-quality meta-research (Saini & Shlonsky 2012). For example, Harris Cooper's (2017) seminal handbook presents a 20-point checklist for his seven-step approach to meta-research; Mieke Heyvaert *et al.* (2017) suggest an eight-step approach, with sub-steps for different parts; Evan Ringquist (2013) has a six-step approach; and Claire Howell Major and Savin-Baden (2010) a ten-step approach. What all these handbooks agree on, however, is that meta-research must "leave a trail for others to follow" by being systematic, transparent, and replicable – and, ideally, updatable over time (Eklund Karlson & Takahashi 2017).

3.2. The first step(s) of meta-research

It is widely acknowledged that a solid piece of meta-research starts with the development of a meta-research protocol. This protocol will guide the researchers during the meta-research process and, more importantly, will help the readers to understand the steps that were taken, how they were taken, and why. As with the stepwise approach discussed earlier, there is no fixed design for a meta-research protocol. There are, however, good examples to follow. For example, protocols such as AMSTAR 2 (A MeaSurement Tool to Assess systematic Reviews, 2nd version), MARS (Meta-Analysis Reporting Standards), PRISMA-P (Preferred Reporting Items for Systematic Meta-Analysis Protocols), and RAMESES (Realist And Meta-narrative Evidence Syntheses: Evolving Standards) have been developed by collaboratives of scholars and are widely accepted as meeting best practice standards

(Gough *et al.* 2012; Cooper 2017; Gurevitch *et al.* 2018). In addition, journals such as *Systematic Reviews* publish protocols for a variety of research syntheses.

Typically, the protocol provides an explanation of the aim or rationale for the meta-research, together with administrative information (including amendments made to earlier iterations of the protocol, information about in-kind or financial support, and the role of the sponsor of the meta-research – if applicable). Good practice requires that the protocol pays detailed attention to the sourcing of studies (such as eligibility criteria, inclusion and exclusion criteria, and the data management system used). Other items include data abstraction, analysis, and synthesis, and an explanation of how the quality and possible bias of studies is assessed. Good practice expects scholars to register or otherwise publish this research protocol "to demonstrate to future readers whether the questions and methods used were decided a priori" (Mayo-Wilson & Grant 2019, p. 473). For example, the International Prospective Register of Systematic Reviews (PROSPERO) was launched in 2011 to allow for the registration of meta-research protocols.[2] In sum, drafting a protocol along these lines forces scholars to think carefully about how they will undertake the meta-research, and what trail they will leave for others to follow.

3.3. Sourcing, inclusion, and exclusion of studies

Typically, an iterative process is followed, in which an initially broad set of possible studies for analysis is reduced to those that will provide the most valid answer to the meta-research question(s) posed. Different sourcing strategies can be used, including exhaustive, selective, and purposeful sourcing. An exhaustive strategy "aims to collect *all* relevant studies", a selective strategy "aims to identify *all* relevant studies within specified limits" (e.g. from a specific policy area or only peer-reviewed publications), and a purposeful strategy "aims to find information-rich studies" (e.g. by focusing on the research of a specific group of scholars) (Heyvaert *et al.* 2017, p. 76). There is increasing recognition that even the most exhaustive approach to sourcing studies cannot identify the full universe of studies. The advice is therefore to aim for "fitness for purpose" and find a balance between rigor and relevance (Booth *et al.* 2016, p. 21). An initial set of studies can be sourced through, for example, keyword searches in databases such as Google Scholar, WorldCat, and Web of Science. Good practice guidelines suggest using multiple databases to collect the initial set of sources (Cooper 2017).

A constant point of debate in the methodological literature concerns the use in meta-research of unpublished or grey literature (Ringquist 2013; Eklund Karlson & Takahashi 2017). The body of peer-reviewed research tends to be biased toward studies with new, sweeping, or otherwise significant findings. Studies that result in repeated, modest, or nonsignificant findings are less likely to be published in academic journals (Franco *et al.* 2014). Sourcing solely from the peer-reviewed literature may give only a part of the full picture. Including grey literature and unpublished literature in the meta-research is suggested as a potential solution to this problem, albeit one that will make the sourcing, inclusion, and exclusion process even more time- and resource-intensive (Mahood *et al.* 2014).[3] Likewise, scholars are warned to not skim over challenges such as dealing with multiple publications on the same study, or synthesizing across studies that report on different qualities or significances of observations (Borenstein *et al.* 2009; Ringquist 2013). Sometimes it will make sense to include multiple publications of the same study, for example when the publications report on truly different findings. Otherwise, it may bias the synthesis, for example when the multiple publications from the same study effectively report the same findings (which would overrepresent them in the full set of findings).

Once the initial set of sources, which may be tens of thousands of publications, has been collected, the process of inclusion and exclusion begins. Typically, the publications are exposed to different rounds of ever more detailed inclusion and exclusion criteria. For example, in a first round all abstracts of the selected publications may be read to assess whether the publication engages with the phenomena of interest. This first round leans, ideally, on the side of caution and is liberal in including publications for the next round. In a second round, parts (for instance the theoretical discussion or methods section) of the publications that have made it through the first round may be screened for the repeated use of keywords (arguably, a low use of keywords may indicate that the publication is not explicitly dealing with the topic of the meta-research). This process continues until those publications most relevant for the meta-research are uncovered, and it is not uncommon to see four or more rounds of inclusion and exclusion.

Good practice requires that at least two individuals assess *all* publications against the inclusion criteria in *each* round (Gough *et al.* 2012; Stoll *et al.* 2019). Good practice also expects that insight is given into the (dis)agreement between the assessors by reporting the agreement percentage or Cohen's kappa (a measure that indicates agreement beyond chance), and preferably both of these (Ringquist 2013; Heyvaert *et al.* 2017). It goes without saying that sourcing, including, and excluding studies is among the most time- and resource-intensive tasks of meta-research. It is therefore encouraging to see the development of software and machine learning tools that aid this process (Gough *et al.* 2012).

3.4. Quality assessment, evidence map, and gap analysis

Once the final set of publications for the meta-research has been selected, these publications are exposed to some form of quality assessment (also referred to as a critical appraisal). This step has long been debated in the methodological literature (Carroll & Booth 2015; Valentine 2019). The aim of the quality assessment is to evaluate "whether a primary research study has properly been conducted and whether the evidence it produced can be trusted" (Heyvaert *et al.* 2017, p. 114). This quality assessment is explicitly *not* meant to repeat the peer review process (and make a judgment about the publication), but is meant to gain insight into the risks associated with a publication's validity, reproducibility, and value (Petticrew & Roberts 2006; Saini & Shlonsky 2012). Helpful tools, such as the Qualitative Research Quality Checklist and the Critical Appraisal Skills Programme (CASP) checklists, are available to guide this process. The quality assessment is, ideally, included in the meta-research publication, to give readers an insight into the overall quality of the source material, and it can be used to trim down the set of publications for the meta-research further (e.g. by excluding high-risk publications).

Second, there is an opportunity at this point to build a map of the available evidence uncovered and, as a related task, to undertake a gap analysis of (possible) blind spots in that knowledge base (Gough *et al.* 2012). An evidence map, for example one in tabular form, can give an overview of the main geographical locations of the (first) authors of publications, the main geographical locations of the empirical research, the main groups of scholars, the main theoretical frameworks or research designs applied, and so on. Building such a map logically also indicates the gaps and blind spots in the set of publications distilled for the meta-research. The evidence map and the gap analysis may result in a (slight) refocus of the meta-research or a broadening of the source material (which implies going through the above steps again with a complementary set of publications), or both. Whatever the case, at this point good practice expects that scholars undertaking meta-research discuss the extent of the missing data or gaps in the evidence base, and how this may affect the findings of the meta-research (Cooper 2017).

3.5. Analyzing publications

Having done (and left a trail of) the above, the final set of publications included for the meta-research can be analyzed. In effect, at this point meta-research becomes similar to original research, and conventional methods for quantitative, qualitative, and mixed-method research can be applied (Cooper 2017; Heyvaert *et al.* 2017). The main difference is that in original research we expose data collected from databases, surveys, interviews, case studies, and so on to accepted tools and logic, and have to meet accepted reporting criteria, whereas in meta-research we expose the findings from original research to such tools and logic (Vevea *et al.* 2019; Wilson 2019).

The methodological literature presents a wide variety of techniques and approaches for analyzing publications and presenting results (Card 2015; Booth *et al.* 2016; Cooper *et al.* 2019). These range from "traditional" statistical average effect size techniques (including logic models and network analyses) to more recent integrative approaches that contribute to explanation building (including Bayesian meta-analysis and realist synthesis). It is beyond the aim (and word limit) of this article to discuss these various techniques and approaches in detail. Readers interested in carrying out meta-research are strongly recommended to explore the techniques and assess which is most suitable for their project – the various handbooks identified in Table 1 are a good starting point.

4. An illustrative application: Meta-research of empirical responsive regulation studies

To illustrate and gain a better understanding of the value of meta-research for regulation and governance scholarship, a realist evidence synthesis of the peer-reviewed empirical responsive regulation literature is presented, following the meta-research methodology and good practices discussed above.

Responsive regulation, as conceptualized by Ayres and Braithwaite (1992), is one of the most widely discussed regulatory theories and approaches to regulation in the academic literature. It is also widely followed in regulatory practice around the world (Braithwaite 2011; Parker 2013). In addition, a realist evidence synthesis is chosen here, as this is a promising approach for meta-research on mid-level theories and policy programs, including the type of regulatory interventions, regimes, and systems with which readers of *Regulation & Governance* engage (Booth *et al.* 2016; Heyvaert *et al.* 2017). In a nutshell, a realist evidence synthesis is "a theory-driven synthesis aimed at unpacking mechanisms of how complex programs or interventions work (or why they fail) in particular contexts and settings" (Heyvaert *et al.* 2017, p. 237). It is "used to articulate underlying programme theories and then interrogate existing evidence to find out whether and where these theories are pertinent and productive" (Booth *et al.* 2016, p. 260). It allows for the use of both qualitative and quantitative source material, and for the mapping, exploration, and interrogation of both the performance of a program or theory (i.e. its effectiveness) and the dominant themes discussed in the source material that look beyond mere quantitative performance (i.e. discussions on ethical and epistemic challenges with responsive regulation emerged as a theme in the responsive regulation literature and are, therefore, addressed in what follows).

4.1. A very brief summary of responsive regulation

It will be safe to assume that most readers of *Regulation & Governance* are somewhat familiar with responsive regulation theory and the book from 1992 that introduced it: *Responsive Regulation: Transcending the Deregulation Debate*.[4] Here, only the main points that are relevant to an understanding of the evidence synthesis that follows are addressed.[5] In short, Ayres and Braithwaite (1992) argue that strict, government-led, command-and-control regulatory policies are often not the best way to address societal problems, and neither are laissez-faire policies that rely on market competition. They introduce responsive regulation as a general regulatory strategy that seeks to build on the strengths of both these approaches and overcome their weaknesses.[6] Responsive regulation is conceptualized as "an attitude that enables the blossoming of a wide variety of regulatory approaches" and not "a clearly defined program or a set of prescriptions concerning the best way to regulate" (Ayres & Braithwaite 1992, p. 5). Elsewhere, Braithwaite explains that the basic idea of responsive regulation "is that governments should be responsive to the conduct of those they seek to regulate in deciding whether a more or less interventionist response is needed" (Braithwaite 2002, p. 29). Responsive regulation assumes a general willingness of citizens and firms to regulate themselves. The challenge is to find those regulatory interventions that provide the best mix of government intervention and self-regulation. Ayres and Braithwaite (1992) argue that it is possible to meet this challenge. Governments must have in place a set of escalating forms of government intervention that come into force if delegated forms of market regulation or other less intrusive ways of government regulation do not yield desirable outcomes.

Responsive regulation is probably best known for the different "enforcement pyramids" it introduces. One of these illustrates the targeting of individual citizens or firms. It shows how a regulator can engage with a citizen or firm through a set of escalating regulatory interventions (*tit-for-tat*): from explaining the purpose of the regulation to seek compliance (persuasion), via warning letters or civil penalties (deterrence), to criminal penalties or license revocations (the full force of the law). The enforcement pyramid is a heuristic that seeks to illustrate how compliance is likely to be achieved through "soft" regulatory interventions, so long as the regulator can escalate to a severe response *and* is willing to use the most critical intervention. Other suggested approaches are tripartism, enforced self-regulation, and partial-industry regulation.[7] *Tripartism* introduces the idea of empowering citizens' associations, for example, to overcome the risk of capture and corruption in the traditional two-way interactions between government and a regulated industry. *Enforced self-regulation* requires firms to write their own sets of corporate rules, which are then publicly ratified and enforced. *Partial-industry regulation* seeks to gain leverage from the competitive conduct of an entire industry by regulating some firms but not others.

4.2. Aim of the evidence synthesis of empirical responsive regulation literature

When *Responsive Regulation* was published in the early 1990s, it filled a gap by suggesting a "third way" between intrusive government regulation and laissez-faire policies. Over time, that specific aspect of the book seems to have become somewhat lost. In engaging with responsive regulation, most scholars, policymakers and regulatory practitioners focus on enforcement pyramids (tit-for-tat), and often take these as a prescription for how regulators should work (Mascini 2013; Parker 2013). In addition, little empirical work appears to be available that gives insight into the opportunities and constraints of responsive regulation in the day-to-day practice of regulation (Braithwaite 2002; Lehmann Nielsen & Parker 2009; Simpson 2013; Zhu & Chertow 2019). In short, it is still not known *whether (on average) responsive regulation outperforms alternative regulatory strategies*; *under what circumstances responsive regulation works best*; or *what it is about responsive regulation that makes it work better (or worse) than alternative regulatory strategies, and for whom, where, in what respects, and why*. Those three questions are central to the evidence synthesis that follows.

Of course, the 1992 book is "but" a moment in time for responsive regulation as a theory. The theory has developed since (and was "under construction" before), in part because of some of the empirical material that is the basis of the evidence synthesis that follows. Logically, the starting point of an evidence synthesis of a "living" theory may affect the findings. That being said, the central premises of the 1992 book recur in later iterations of the theory (e.g. Braithwaite 2002, 2011; Braithwaite *et al.* 2007), which reduces the challenges that come when empirical research that builds on later iterations is assessed against the original 1992 premises.

4.3. Contribution to the broader literature on responsive regulation

The evidence synthesis that follows, and the leading questions in particular, move well beyond the critique that Braithwaite has made elsewhere of (earlier) reviews of the responsive regulation literature (Braithwaite 2016a). Rather than merely presenting a reductionist or positivist answer to a quantitative performance question, the realist evidence synthesis that follows aims to understand how responsive regulation performs in both qualitative and quantitative terms (Pawson 2002; Heyvaert *et al.* 2017). In this way, the evidence synthesis presented here is more in line with Braithwaite's own attempts to synthesize the evidence on responsive regulation (Braithwaite 2016b), albeit in a more transparent, replicable, and systematic manner.

In addition, building on the logic (and opportunities) of realist evidence synthesis, the evidence synthesis that follows aims to understand the documented differences in qualitative and quantitative performance. Besides reporting on how responsive regulation performs in day-to-day practice, two themes recur in the literature: how to make responsive regulation perform better, and the ethical and epistemic challenges that come with this approach to regulatory governance and regulatory delivery.

4.4. Leaving a trail for others to follow

For the evidence synthesis, a protocol was developed and followed for sourcing and coding documents. This protocol builds on four (partly overlapping) tools, standards, and protocols that are conventional for the type of evidence research presented here (including AMSTAR 2, MARS, and PRISMA-P that are discussed above). The protocol was developed and published online before the sourcing of studies began. The protocol was updated twice during the round of including and excluding studies and processing data, and the updated protocols were published online accordingly; the final protocol is available as Appendix S2.

4.5. Sourcing, inclusion, and exclusion of studies

It is a daunting task to draw a comprehensive sample of empirical responsive regulation literature that is at the same time practical. The book *Responsive Regulation* has been cited more than 5,000 times since its publication.[8] In selecting the source material from this "complete universe," no restrictions were set by time, length, or repetitions of the studies that were to be included, and nor were there any restrictions on the setting(s) of studies; however, only articles written in English were included. Also, only published peer-reviewed articles, including "online first" and "early access" publications, were included. The articles included in the evidence synthesis were systematically sourced from different databases: WorldCat, Scopus, and Web of Science. In the first step, these databases were explored to identify articles that were likely to engage with Ayres and Braithwaite's ideas on responsive regulation:

- From WorldCat, all articles with the words "responsive regulation" in their keywords, published in English since 1992 in the subject areas "business and economics," "law," "sociology," and "political science," were included. This search resulted in 132 documents.
- From Scopus, all articles with the words "responsive regulation" in their titles, abstracts, or keywords, published in English since 1992 in the subject areas "social science," "business, management and accounting," and "economics, econometrics and finance." were included. This search resulted in 105 documents.
- From Web of Science, all articles with the words "responsive regulation" in any searchable field and all articles citing the book *Responsive Regulation* using the "cited reference search" tool for the combination ayr* (cited author) AND res* (cited work), published in English since 1992 in the disciplines "business," "criminology," "law," "economics," "management," "political science," "public administration," "public policy," "social sciences interdisciplinary," and "sociology," were included. This combination of search terms for the cited reference search is likely to overcome the most common typographical errors in citations of the book. These searches resulted in 94 and 807 documents, respectively.

This set of documents was supplemented with all articles from three special issues on responsive regulation: volume 7, issue 1 of *Regulation & Governance* (2013); volume 44, issue 3 of the *University of British Columbia Law Review* (2011); and volume 29, issue 1 of *Law and Policy* (2007). In addition, all peer-reviewed English language articles from the "narrative review" sections of earlier publications that sought to draw general lessons from empirical research on responsive regulation were included (Baldwin & Black 2008, pp. 62–64; Braithwaite 2016b, pp. 8–18; Mascini 2013, pp. 50–53; Simpson 2013, pp. 324–325; Walker-Munro 2019, p. 118). This sourcing process, akin to "reference list checking" (Gough *et al.* 2012, p. 114), resulted in 86 documents. After removing duplicates (n = 212), this initial search resulted in a set of 1,012 original peer-reviewed journal articles.

In the second step, article titles, abstracts, and keywords were screened to exclude articles that were unlikely to deal with a responsive regulation approach or that were explicitly not empirical, or both. For this, the following scores were used: "yes" (include), "no" (exclude), "unsure" (include). The screening was carried out by two coders – the author of this article and an experienced, external, paid researcher (a librarian by training). To ensure consistency in how the coders understood responsive regulation, a central terminology for each step of the inclusion and exclusion process was developed and discussed – see Appendix S3. Of the full set of set of 1,012 articles, three articles were not accessible. Of the remaining 1,009 articles, 586 were included in the next round. The agreement percentage between the two coders was 75%, with a Cohen's kappa of 0.51 (representing fair agreement, Heyvaert *et al.* 2017). In this step, a liberal approach to inclusion was taken: if at least one of the coders used the score "yes" or "unsure," the article was included for screening in the next step. Thus, only the combinations of "no" and "no" were excluded in this step.

In the third step, the research design section (or similar section) of each article was read to exclude articles that did not deal with a responsive regulation approach, and the method section (or similar section) of each article was read to exclude articles that were explicitly not empirical. For this, the following scores were used: "yes" (include), "no" (exclude), "unsure" (include). Of the 586 articles, 123 were taken forward to the next step. The agreement percentage between the two coders was 95%, with a Cohen's kappa of 0.82 (representing excellent agreement, Heyvaert *et al.* 2017). In this step, all discrepancies for which one coder gave the score "yes" and the other "no" were resolved; but, as in the previous step, articles that were coded with a combination of "yes" and "unsure," or "unsure" and "unsure," were taken forward to the next step.

In the fourth step, the articles were read in full to exclude articles that did not deal with a responsive regulation approach or that were explicitly not empirical, or both. To come to a final decision on articles to include in the review, all disagreements were discussed and resolved. From this step, 65 articles were included for further analysis: 30 articles were identified as providing direct insights into the performance of responsive regulation (mainly in terms of its effectiveness or efficiency), and 35 articles were identified as providing relevant insights into the working of responsive regulation while not explicitly evaluating its performance.[9] The latter group of articles typically discuss how responsive regulation could be improved, or why it is difficult to assess the direct

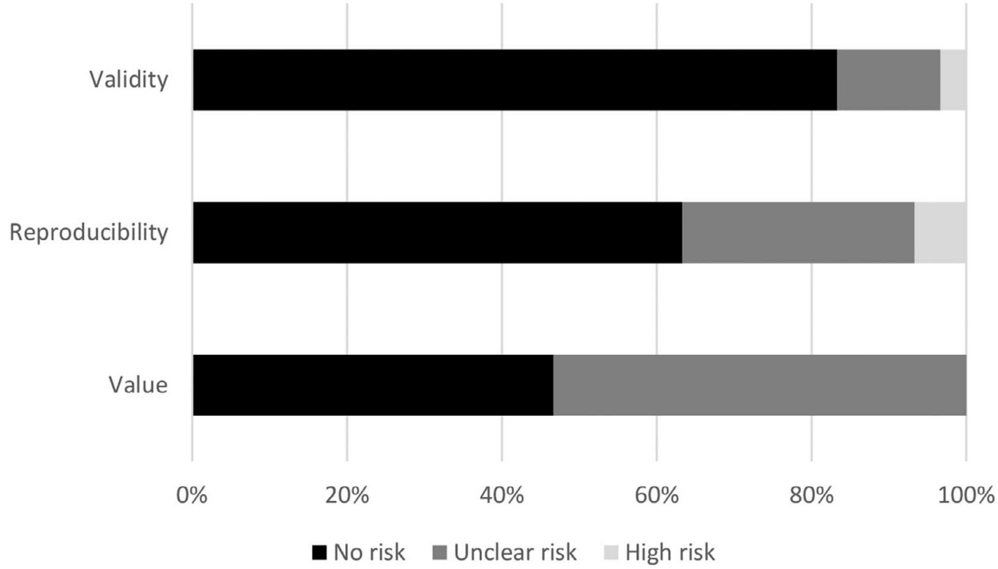

Figure 1 Quality assessment of articles that provided direct insights into the performance of responsive regulation (n = 30).

performance of responsive regulation. In a final step, the 30 articles providing direct insights were assessed to cluster together those that reported on the same study (to prevent the "double counting" of individual studies or study settings).[10] In this step, 24 unique study settings were observed.

This full process took approximately 300 hours between the two coders over a 2-month period. Appendix S4 gives a flowchart of this process.

4.6. Quality assessment

The quality of the 30 articles that provide direct insights into the performance of responsive regulation was assessed using the CASP checklists. Appendix S5 gives an overview of the checklist used. Figure 1 gives an overview of the findings.

Figure 1 indicates that a large majority of the studies included do not indicate a risk in terms of their validity (n = 25). These studies have clear research designs, use appropriate methods to answer the questions they pose, and use sufficiently rigorous approaches to data collection and analysis. Most of them do not indicate a risk in terms of their reproducibility (n = 19). These studies provide a transparent and (likely) replicable data analysis, and a clear statement of findings. Finally, just under half of the studies included do not indicate a risk in terms of their value (n = 14). These studies make clear statements of their aims, answer the question they pose initially, and provide a reflection on their limitations. Overall, this quality assessment does not raise red flags. The overall risk of the full set of studies included is modest, at best, and the scores for "value" are mainly explained by an absence of reflections on limitations in many of the studies included.

4.7. Evidence map

From the 65 articles, data were abstracted following the PICO criteria (participants, interventions, comparators, outcomes)[11]:

- *Study designs*: Of the 30 articles that provided direct insights into the performance of responsive regulation, 13 articles used qualitative data, 14 quantitative data, and the remaining three a mixed set of data. Of the 35 articles that provided additional insights into the performance of responsive regulation, 15 used qualitative data, 19 quantitative data, and the remaining one a mixed set of data.
- *Participants*: Most studies come from western countries (n = 62), with Australia (n = 20) and the United States (n = 16) being the dominant research settings. The articles addressed a range of policy areas, with none being dominant. These areas included: business regulation, consumer protection and public safety, environmental sustainability, finance, operational health and safety, and taxation.

- *Interventions*: In the 30 articles that provided direct insights into the performance of responsive regulation, there were 23 examples of a tit-for-tat or pyramidal approach to regulation, seven of enforced self-regulation, three of tripartism and one of restorative justice (some articles studied more than one form of responsive regulation, which explains why the total is larger than 30). In the 35 articles that provided additional insights, there were 18 examples of a tit-for-tat or pyramidal approach to regulation, two of tripartism, one of restorative justice, and 14 situations where an "ideal type" form of responsive regulation was contrasted with the empirically observed setting.
- *Comparators*: Most articles (n = 39) directly compared a current responsive regulation situation with an earlier or external situation (e.g. traditional government-led command-and-control regulation, or laissez-faire market competition). The remaining articles (n = 26) explained the effects of responsive regulation without making such comparisons.
- *Outcomes*: The following outcome indicators were set (a priori) to evaluate the observed performance of responsive regulation across the set of 30 articles that provided direct insights: changes in compliance level, compliance costs (for the regulator and regulatee), and compliance achievement time (for the regulator and regulatee). These indicators were scored as "positive," "no effect," and "negative" when explicit data on effects were provided; or, as "it depends" when the observed effect was explained in qualified and context-specific terms. The research was open to uncovering additional indicators, including improved relationships between regulators and regulatees, spill-over effects (e.g. higher levels of compliance because regulatees became aware of the new approach to regulation), and improved working conditions for regulatory staff. The findings are presented in what follows.

4.8. Gap analysis

The main risk of bias in the evidence review lies in the fact that only peer-reviewed English language articles were included. As explained above, not all research on responsive regulation makes it through the journal peer review process, and not all research on responsive regulation from around the world is carried out by academics. The exclusion of nonpublished academic work, academic publications other than peer-reviewed articles, and non-academic publications logically limits the number of observed applications of responsive regulation in real-world settings (Mahood *et al.* 2014). Second, there is a risk that positive experiences with responsive regulation are over-represented in the academic literature (a selective reporting bias) and thus in the final set of articles distilled from this literature (Vevea *et al.* 2019) – as becomes clear in what follows, however, only one in three real-world examples discussed in the 30 articles that provide direct insights into the performance of responsive regulation were explicitly evaluated as having a positive effect by the original authors. Unfortunately, because of the types of study designs and types of data used in the included articles (predominantly single-n or small-n studies), it was not possible to run formal tests for sample biases (c.f. Hardwicke *et al.* 2020).

Perhaps more problematic than this possible publication bias is the limited scope of the evaluative empirical research identified. Most of the studies provided insight into the application of responsive regulation in western countries (95%) and were carried out by scholars from western countries (95%). This leaves the global south largely uncovered in this evidence synthesis. Of the geographic areas covered, 31% of the studies presented evidence from Australia and 25% from the United States; likewise, an Australian-based researcher is the first author of 17% of the publications, and a researcher based in the United States the first author of 15%. This makes these two countries vastly overrepresented in this evidence synthesis. It should be noted, however, that 15% of the articles reported on data obtained in non-English speaking countries, and 15% of the articles were written by scholars from non-English speaking countries. Finally, the tit-for-tat or pyramidal approach to regulation represents 68% of all the real-world cases studied; tripartism and restorative justice were both represented in fewer than 10%; and partial-industry regulation in none. This confirms earlier concerns that scholars are drawn toward some parts of the responsive regulation theory, but not toward others.

4.9. Analyzing publications

To gain a rich understanding of the performance of responsive regulation, and particularly an understanding of the circumstances under which responsive regulation was found to perform best, all 65 articles were read closely and coded in Atlas.ti (a computer program that supports the systematic analysis of complex data) using the set of codes presented in Appendix S6. These coded data were explored and analyzed using Atlas.ti. Using this approach, the articles were systematically explored, and insight was gained into the "repetitiveness" and "rarity" of the insights reported. The findings of this analysis are presented in the sections that follow.

5. Findings

Since the publication of their book, Ayres and Braithwaite's ideas on responsive regulation have been applied in a wide variety of policy areas and in a wide variety of countries (Woods *et al.* 2010; Ivec *et al.* 2015). Overall, scholars and practitioners appear to be supportive of the normative arguments (and the heuristics and strategies) presented in *Responsive Regulation*, but some of them question whether it is possible to apply these in regulatory practice. Also, it is not always clear from the documented examples whether the regulators have purposefully followed a responsive regulation heuristic or strategy (along the lines of Ayres and Braithwaite's ideas), or whether those documenting the examples have forced them into a responsive regulation mold. In addition, scholars have mapped examples of alternative applications of responsive regulation. For example, Zhu and Chertow (2019) observed that local regulators in Jiangsu Province (China) set out to enforce a newly introduced energy efficiency regulation in a formal and coercive manner before switching to a more accommodative manner. This "inverse" enforcement pyramid illustrates that the responsive regulation theory is somewhat pliable, and that some care is warranted when drawing lessons from such studies.

5.1. At first glance: It depends

When considering the 24 unique examples of responsive regulation discussed in the set of 30 articles (see above), responsive regulation was found to outperform alternative strategies in eight cases and to perform worse than the alternatives in six cases. In addition, in ten cases the authors reported that the performance of responsive regulation was unclear: it was highly dependent on how the responsive regulation was implemented by (individual) regulators, how it was experienced by (individual) targets of regulation, and how it interacted with other regulatory reforms that were implemented in parallel. In these cases, the authors tended to stay away from making a judgment about whether responsive regulation performs better or worse than alternative strategies.

The positive performance reported ranged from being substantial (e.g. increased tax revenue for the Florida Department of Revenue (Christian 2017); increased energy savings by Chinese firms (Zhu & Chertow 2019) to being only modestly better (e.g. increased compliance with human rights standards by global garment manufacturers (Islam & McPhail 2011). However, the increased performance was often not quantified (this is the case in half of the positive performance observations[12]). This also holds for cases with negative or unclear performance: the observed performance was only reported in qualitative terms. This finding is, of course, not surprising. Scholars of regulation – and governments – often find it very difficult to measure the exact performance of a regulatory strategy, which makes one-on-one quantitative comparisons of regulatory strategies virtually impossible. In addition, many studies were carried out only a few years (typically fewer than five) after responsive regulation was introduced. It could be that a formal change of a regulatory regime needs more time to settle before its anticipated results are achieved.

To cut a long story short, from an overview of the evidence base, the answer to the question of whether responsive regulation (on average) outperforms alternative regulatory strategies is best summarized as: "it depends." Given that it is close to 30 years since the publication of *Responsive Regulation*, and given the ongoing (normative) arguments for and against the use of responsive regulation in practice, this is anything but a satisfactory answer. Looking beyond these direct performance observations, however, the full set of 65 articles provides a wealth of insights about the factors that affect the performance of responsive regulation.

5.2. At second glance: We know a great deal about how to make responsive regulation work well

While the evidence base falls short of making a one-size-fits-all argument for or against responsive regulation, it does provide us with a wealth of information about how to best implement responsive regulation – or, at least, it raises a set of questions about how to implement it better.

5.2.1. Is training of staff in responsive regulation essential?

A recurring observation is that responsive regulation requires careful implementation. Street-level bureaucrats, in particular, need guidance documents, training, or both, to be able to deal with the increase in discretionary space that comes with responsive regulation, and specifically with the tit-for-tat strategy (Blanc 2018). If this is not done, there is a risk of inconsistency in enforcement (Mascini & van Wijk 2009), favoritism (Zhu & Chertow 2019), abuse of process (Lewis *et al.* 2017), uneven application of the law (O'Neill 2018), or rule ambiguity (Quilter & McNamara 2015). Inconsistent enforcement by and between street-level bureaucrats may undermine the knowledge of the targets of regulation (May & Wood 2003). An innovative approach that has been identified as a way to increase consistency across street-level bureaucrats is the introduction of a peer review system in which bureaucrats are asked to reflect on each other's and their own enforcement practices. There is evidence that such a system makes the responsive regulation enforcement process less arbitrary, more transparent, and more accountable (Ho 2017).

These findings hold not only at the level of street-level bureaucrats but also at organizational and supra-organizational levels: differences in the application of responsive regulation between regulatory agencies within a country, or even between countries as a whole, may undermine the performance of regulatory regimes (Sampson & Bloor 2007). In sum, in applying regulatory pyramids, the main challenge is the shifting between different regulatory actors: "Attractive as the theory of responsive regulation is at the primary, interactional level … it does not get us as far when we consider the effects of shared and interlocking jurisdictions in multi-level regulatory systems. The elements of responsiveness, difficult to produce at the primary level [street level], are even more difficult to produce at secondary [entities and activities that create and maintain infrastructure (the rating and certification bodies) and entities such as courts that hear appeals] and tertiary [policy making] levels" (Heimer 2011, p. 691) (see also Lehmann Nielsen & Parker 2009).

5.2.2. Is it essential to explain to regulatory targets what responsiveness means?

A second recurring observation considers the transparency and clarity of the responsive regulation process and the regulatory dialogues involved. Put simply, improved or simply more communication between regulators and their targets is by no means a guarantee of enhanced compliance (Lehmann Nielsen & Parker 2009; Curtis *et al.* 2016; Eungkyoon *et al.* 2016; O'Neill 2018). Targets of regulation do not always understand or appreciate that the purpose of a responsive regulation approach is to seek compliance in collaboration with the regulators. For example, those who are less knowledgeable about the regulations to which they are subject are likely to regard escalation up a regulatory pyramid as threatening or even bullying (Winter & May 2001), or, as a communication process, responsive regulation may simply be too demanding for (some) targets (Tränkle 2007). Targets are also likely to misinterpret the regulators and to consider them, in general, to be harsher than the regulators intend (Mascini & van Wijk 2009).

Responsive regulation is expected to be more effective if the regulated sector knows that the regulator will follow through on its escalation strategy (Short & Toffel 2010; Toffel & Short 2011); likewise, if the regulator is known to be "soft," this is likely to undermine overall compliance in the industry (van Erp 2011). In a nutshell, it may be helpful for a regulator to explain the responsive regulation strategy to its targets at the start of the regulatory relationship, and to have in place transparent public procedures and other measures that help to prevent the relationship between regulator and target from becoming too intimate (de Boer & Eshuis 2018). For example, when bringing third parties into a regulatory regime, regulators may have to explain exactly why it is in their interests to be involved (Kingsford Smith 2011).

5.2.3. Is the timing of the start and escalation of the responsive regulation dialogue important?

A third recurring observation builds on these insights. Responsive regulation is expected to be more effective when regulators and their targets have multiple interactions (rather than a one-off engagement), and when the targets of regulation have (or are reminded of) a long-term time horizon (van Duin *et al.* 2018). The start of a responsive regulatory dialogue may set the tone for these interactions, and sometimes a penalty or fine at the start of this dialogue may have significant symbolic value, showing that the regulator is severe and has the power to *enforce* compliance (Pires 2008). Also, without explaining what a responsive process entails, there is a risk of a "vicious circle of antagonism" if regulators follow responsive regulation to the letter: harsh enforcement of regulations against already skeptical targets of regulation is likely to make those targets even more skeptical (Mascini & van Wijk 2009).

Indeed, the escalation to a more formal stance can easily "backfire once a certain threshold of formalism has been reached" (May & Winter 1999, p. 627). Besides, in contexts where the rule of law is weak, targets of regulation may be regulator-driven rather than regulation-driven. In such contexts, the "more reasonable and regulatee-friendly approach" suggested by responsive regulation may backfire, as the targets of regulation may see the regulator as weak and easy to manipulate (Yee *et al.* 2016, p. 109). In such contexts, regulators may need to "go hard, go early" when escalating to more intrusive responses. Last but not least, responsive regulation may yield different results across age and socio-demographic groups (Muehlbacher *et al.* 2011).

5.2.4. Is responsive regulation without risks?

A final set of recurring observations highlights the risks of responsive regulation. When the focus of a regulatory strategy is escalation, regulators may choose to pursue the "easy cases" and avoid the "hard" ones (O'Neill 2018). This could simply be because less effort is needed to enforce the easy cases (more bang for one's buck), but it may also be because there are legitimacy and accountability risks in more complex cases. That is, the regulators know that they are more likely to slip up in such cases (Parker 2006). Also, regulators may face severe political or societal criticism if an incident is linked to a "soft and facilitative" approach to regulatory enforcement. They may then be expected to move to harder and more formalistic enforcement. Such a change in regulatory practice may result in a rapid breakdown of trust between regulators and their targets (Gunningham & Sinclair 2009).

Another risk of responsive regulation (particularly enforced self-regulation and other forms of delegated regulation) is that targets of regulation have in place the systems, rules and procedures demanded by regulatory law but do not apply them, or do not have staff who understand how and why to apply them (Hutter 2001). In such situations, it is not enforcement (the use of the benign big gun) but monitoring that keeps the targets of regulation on their toes (Short & Toffel 2010). Finally, tripartism may backfire when third parties are biased toward reporting some violations and not others. For third parties, highly visible but low-risk violations are likely to be easier to observe than less visible but high-risk violations. As a result, a regulatory agency may end up in a situation where it has to process (more) low-risk violations at the expense of processing high-risk violations (Lochner & Cain 1999; Lochner *et al.* 2008).

5.3. At third glance: Ethical and epistemic challenges

Another set of insights that emerges from the full set of 65 articles relates to the ethical and epistemic challenges that come with responsive regulation. In part, these insights explain both the variance in performance observations and the, overall, small number of articles that explicitly assess the performance of responsive regulation. Most of the ethical challenges discussed in the literature can be summarized as: "To be responsive to one stakeholder is to be potentially unresponsive to another" (Bryer 2007, p. 482) – but more is at stake. Many of the epistemic challenges discussed in the literature address the limits of and differences in knowing what it means to be responsive (Baldwin & Black 2008; Lehmann Nielsen & Parker 2009; McKay *et al.* 2015; Sheehy & Feaver 2015). Such challenges are identified at the level of both individuals (actor-centered responsiveness) and regulatory systems (institution-centered responsiveness).

5.3.1. Ethical challenges

A cluster of ethical issues concerns inconsistencies in the application of responsive regulation. From a discretionary ethics point of view (Bryer 2007), such inconsistencies are particularly problematic because the way in which regulators respond to a target's behavior is often largely outside the control of that target (von Holderstein Holterman 2009). Besides, external parties may seek to play the ambiguities of responsive regulation and to influence the relationship between regulators and their targets (Simpson 2013). This introduces a risk that the response of the regulator is influenced, for example, by external societal or political pressures and not by the behavior of the target (Parker 2006; Baldwin & Black 2008; Almond & Esbester 2018).

From a control-centered ethics point of view (responsiveness to the institutional setting within which responsive regulation operates, see further Bryer (2007), responsive regulation may be problematic because regulators may too quickly embrace it as a rational, utilitarian tool (Baldwin & Black 2008). The presentation of a tit-for-tat strategy and the related regulatory pyramids, in particular, builds on the idea that targets of regulation will respond "rationally" (and predictably) to the escalation or de-escalation of the regulator (von Holderstein Holterman 2009). Likewise, responsive regulation assumes that regulators will respond "rationally" (and predictably) to their targets. However, such assumptions of rationality have long been debated (Kahneman 2011). In a similar vein, while responsive regulation aims to achieve compliance through nondomination, it leaves in place a power imbalance: the regulator is always more potent than its target, and can always fall back on using the benign big gun (Daniels 2012; O'Neill 2018).[13] A related control-centered ethical issue is the downshifting of responsibilities from the top of regulatory agencies to frontline workers by expanding their discretionary space (Waller 2007).

A final ethical issue that is argued by some scholars and is worth mentioning here is that responsive regulation plays a high-stake game with trust (which aligns with the notion of deliberative ethics, see Bryer [2007]). Responsive regulation may crowd out trust throughout the regulatory system. If the targets of regulation lose trust in the regulator early on in the regulatory dialogue, when the regulator takes a facilitative stance, they are unlikely to trust the regulator later in the regulatory dialogue once the regulator has escalated to a more formal position (Benkler 2012). Likewise, if a regulator loses trust in its target early on in the regulatory dialogue when seeking compliance through collaboration, the regulator is unlikely to trust the target later on in the dialogue when the regulator has escalated to a more formal and stricter interaction. In a somewhat related manner, punishing a target with "soft instruments" may have unintended spill-over effects. For example, the staff at all levels of a "named and shamed" firm may also feel stigmatized as violators, which is less likely to happen with "traditional" punishments such as a fine imposed on the firm (which is likely only to affect staff in the higher echelons of the firm – if it affects any of them). Such staff may be less trusting of the regulator in a future interaction (Parker 2006; van Erp 2011).

5.3.2. Epistemic challenges

The epistemic challenges uncovered relate to the difficulties of having objective knowledge about a responsive regulator – for a regulatory practitioner, a target of regulation, and a regulatory scholar. For example, Lehmann Nielsen (2006) observed five different types of (actor-centered) responsiveness in inspectors. The first is short-term responsiveness: when observing a violation, inspectors base their response on the gravity and number of violations discovered at the same time. The second is long-term responsiveness: the average severity and the number of previous breaches also affect this response. The third is attitude responsiveness: besides these objective markers, inspectors are influenced by how willing they think their target is to improve its behavior. The fourth is dialogic responsiveness: inspectors' responses are affected by the average number of engagements they have with their targets every year. The fifth and final type is subjective performance responsiveness: inspectors are swayed by their overall, but often individual, assessment of their target's standard of compliance. These different types of responsiveness amplify the inconsistencies in the implementation of responsive regulation that were observed in other studies (Lehmann Nielsen 2006). Insights like these only stress the need for training and institutional guidance when embracing responsive regulation as a heuristic (or philosophy) in regulatory practice.

It is also relevant to recall that responsive regulation, as conceptualized by Ayres and Braithwaite, builds on a republican political philosophy, or even a republican ideal. This ideal centers on citizen participation, a desire of those in power to reduce the domination of those not in power and to encourage each target of regulation to be "a responsible citizen, to be law abiding" (Ayres & Braithwaite 1992, p. 17; Braithwaite 2013; also, Braithwaite 2017). In addition, responsive regulation is not only about more effective or efficient regulatory governance and practice (i.e. "more bang for one's buck"), but also about increased accountability, transparency, equity, and so on, in the regulatory process. Such competing values make it difficult to come to a final verdict as to whether responsive regulation outperforms other approaches to regulation. After all, such a verdict would be unjustly reductionist and over-simplified.

6. Conclusion

This article has presented a narrative review of selected meta-research methodology literature, before presenting a realist evidence synthesis of the empirical peer-reviewed responsive regulation literature. Its primary aim was to expose readers of *Regulation & Governance* to meta-research as a relevant area of research that may help us, as a community of regulation and governance scholars, to strengthen the overall knowledge base of our field. The narrative review of the methodological literature and the illustrative example of the responsive regulation evidence synthesis may inspire others to follow suit and to approach other areas of regulation and governance knowledge in a similar manner. Here it should again be noted that, whilst being primarily a set of desk-research activities, meta-research is a time- and resource-intensive process and has its place and purpose. Not every research project needs an evidence synthesis of the coverage and coherence presented here. For many research projects, a more modest review of the literature will likely be sufficient – but even in those projects it pays to be systematic and leave a trail for others to follow. There is no fixed way of "doing" meta-research. The good practices allow for a lot of freedom and flexibility. After all, "what we are working with here is a dynamic theory of evidence" (Cooper 2017, p. 319).

The secondary aim of this article was, at this point, to call on the editors of *Regulation & Governance* to take meta-research seriously in the journal. Since the journal was launched, no meta-research proper has been published there (effectively, earlier review articles in the journal are narrative reviews). This is a blind spot in the journal. But one that should not be too difficult to remedy. The editors could, for example, include a meta-research article in each issue of the journal in a dedicated meta-research section, or collate a, say, biannual meta-research special issue. Meta-research could be commissioned from leading scholars in the field; or, perhaps better, early career researchers could be explicitly invited to develop a regulation and governance oriented meta-research publication as part of their (PhD, post-doc, or other) research projects. The editors could require prospective meta-researchers to submit their research protocols for peer review within the *Regulation & Governance* community (before undertaking the meta-research) as one of the conditions for possible later publication. After all, our whole community will be very well served by meta-research projects of this kind as they will save all of us the hassle of ploughing through the ever-growing knowledge base to find promising avenues for future research (and, for the early career researcher, a meta-research article in *Regulation & Governance* will likely become one of their most cited articles in those challenging early years).

The third aim of this article was to listen carefully to Edward Wilson, who was quoted at the very start. That quote continues as follows: "The world henceforth will be run by synthesizers, people able to put together the right information at the right time, think critically about it, and make important choices wisely" (E. Wilson 1999, p. 294). For some time now, we have seen regulators struggling with responsive regulation – because with all the writing on responsive regulation they cannot see the wood for the trees. For them, this article puts together the best available information on responsive regulation, and clusters this information in a way that aids critical thinking about it, so that they can make important choices about using (or not using) responsive regulation wisely. To sum up, first, just having responsive regulation strategies in place (e.g. a regulatory pyramid) is no guarantee that these strategies are used in regulatory practice, let alone that they are used well (Welsh 2009; Bird & Gilligan 2016; Hedges *et al.* 2017). Staff need training, targets of regulation need explanation, and agencies need alignment to enable responsive regulation to be operationalized in practice. Second, just cherry-picking responsive elements will not result in a satisfactory regulatory regime (Moss 2019). While *Responsive Regulation* provides hands-on strategies, these need to be implemented within the larger philosophy of the theory – a flexible and context-sensitive approach to regulation that seeks compliance through the least intrusive intervention possible. Third, a static responsive regulation regime may lose its effectiveness over time when staff at regulatory agencies change, or when targets learn to roll with the punches of the regime (Szejnwald Brown *et al.* 2001). Ideally, as with any approach to regulation, a responsive regulation regime needs to involve an element of learning and updating based on lessons learnt (Baldwin & Black 2008). Fourth, because of its flexibility, the language of responsive regulation can easily be abused to push a neo-liberal deregulation agenda (Tombs & Whyte 2013) or a repressive regulation agenda (Yee *et al.* 2016). Regulators and their targets need to keep asking questions about *what* exactly responsive regulation implies, in their setting and in their applications. Fifth, and following on from all of the above, responsive regulation should ideally be introduced as part of a suite of regulatory reforms (Blanc 2018). Without additional staff training, transparent and accountable public procedures, and ongoing learning and improvement, it is unlikely that the real-world application of responsive regulation will live up to its theoretical promises (Lehmann Nielsen 2006).

Endnotes

1 Practically speaking, this section is a narrative review of the literature on meta-research. Appendix S1 provides an overview of the full set of literature sourced for this review, as well as the set of codes used to extract insights from this literature.

2 Some academic journals allow (or even expect) the registration of meta-research protocols. Organizations such as the Cochrane Collaboration and the Campbell Collaboration also allow for the registration of protocols.

3 A helpful source for tracing grey literature is GreyNet, the Grey Literature Network Service.

4 There is a debate as to whether (some) (more) credit for the theory should go to Nonet and Selznick (2009 [1978]) and Silbey (1984), who discussed ideas related to responsive regulation well before Ayres and Braithwaite (1992) published their work. Exploring that debate further is, however, beyond the aim of this article.

5 Readers interested in broader discussions of the theory are strongly advised to read the original book, as well as the various articles published in special issues of academic journals that celebrate the book. These issues are: *Law and Policy*, vol. 19, issue 1 (2007), *University of British Columbia Law Review*, vol. 44, issue 3 (2011), and *Regulation & Governance*, vol. 7, issue 1 (2013).

6 Ideas on responsive regulation were, according to Ayres, largely developed by Braithwaite (Ayres 2013).

7 Braithwaite later introduced *restorative justice* as another approach to responsive regulation (Braithwaite 2002).

8 Source: www.google.scholar.com (2 August 2020).

9 Effectively, it is only the 30 articles that inform the first sub-section of the findings ("At first glance: It depends"), and the full set of 65 articles that inform the second and third sub-sections.

10 In this study, no article was excluded, but if a similar finding was reported multiple times, it was treated as a single observation.

11 As with so many tools in meta-research, the PICO criteria were initially developed in the health sciences, but they have since been embraced in other fields (Cooke *et al.* 2012).

12 The performance of responsive regulation is mainly quantified in studies of tax regulation (Braithwaite 2003; Job *et al.* 2007, see also the wide variety of non-peer-reviewed research into responsive regulation and taxation on www.regnet.anu.edu.au/research/publications). Here it should be noted that taxation is, in general, considered to be a highly efficient way of regulating the public good, and one in which modest efforts made by governments (including a turn to responsive regulation) are likely to yield substantial results (Smith 2015). Looking beyond mere cost-effectiveness, however, other scholars have questioned how well responsive regulation can be applied in this area (Waller 2007; Leviner 2008).

13 As one reviewer noted, it may be that some scholars interpret non-domination in a different way from Ayres and Braithwaite—i.e., the republican theory of non-domination does not require no power imbalance, but rather that power is exercised in accordance with the rule of law.

References

Almond P, Esbester M (2018) Regulatory Inspection and the Changing Legitimacy of Health and Safety. *Regulation & Governance* 12(1), 46–63.
Ayres I (2013) Responsive Regulation: A Co-author's Appreciation. *Regulation & Governance* 7(1), 145–151.
Ayres I, Braithwaite J (1992) *Responsive Regulation. Transcending the Deregulation Debate*. Oxford University Press, New York.
Baldwin R, Black J (2008) Really Responsive Regulation. *Modern Law Review* 71(1), 59–94.
Bearfield DA, Eller WS (2008) Writing a Literature Review: The Art of Scientific Literature. In: Yang K, Miller GJ (eds) *Handbook of Research Methods in Public Administration*, 2nd edn, pp. 61–72. Taylor & Francis Group, Boca Raton, FL.
Benkler Y (2012) Law, Policy and Cooperation. In: Balleisen EJ, Moss D (eds) *Government and Markets: Toward a New Theory of Regulation*, pp. 299–332. Cambridge University Press, Cambridge.
Bird H, Gilligan G (2016) Deterring Corporate Wrongdoing: Penalties, Financial Services Misconduct and the Corporations Act 2001 (Cth). *Company and Securities Journal* 34(5), 332–359.
Blanc F (2018) Tools for Effective Regulation: Is "More" Always "Better"? *European Journal of Risk Regulation* 9(3), 465–482.
Block J, Kuckerts A (2018) Seven Principles of Effective Replication Studies: Strengthening the Evidence Base of Management Research. *Management Review Quarterly* 68(4), 355–359.
de Boer N, Eshuis J (2018) A Street-Level Perspective on Government Transparency and Regulatory Performance: Does Relational Distance Matter? *Public Administration* 96(3), 452–467.
Booth A, Sutton A, Papaioannou D (2016) *Systematic Approaches to a Successful Literature Review*, 2nd edn. Sage, London.
Borenstein M, Hedges L, Higgins J, Rothstein H (2009) *Introduction to Meta-Analysis*. Wiley, Chichester.

Braithwaite J (2002) *Restorative Justice and Responsive Regulation*. Oxford University Press, Oxford.
Braithwaite J (2003) Meta Risk Management and Responsive Regulation for Tax System Integrity. *Law & Policy* 25(1), 1–16.
Braithwaite J (2011) The Essence of Responsive Regulation. *University of British Columbia Law Review* 44(3), 475–520.
Braithwaite J (2013) Relational Republicanism. *Regulation & Governance* 7(1), 124–144.
Braithwaite J (2016a) In Search of Donald Campbell: Mix and Multimethods. *Criminology and Public Policy* 15(2), 2.
Braithwaite J (2016b) *Restorative Justice and Responsive Regulation: The Question of Evidence. Regnet Research Papers 2016 No. 51*. RegNet Australian National University, Canberra.
Braithwaite J (2017) Types of Responsiveness. In: Drahos P (ed) *Regulatory Theory: Foundations and Applications*, pp. 117–132. ANU Press, Canberra.
Braithwaite J, Makkai T, Braithwaite V (2007) *Regulating Aged Care: Ritualism and the New Pyramid*. Edward Elgar, Cheltenham.
Braver SL, Thoemmes F, Rosenthal R (2014) Continuously Cumulating Meta-Analysis and Replicability. *Perspectives on Psychological Science* 9(3), 333–342.
Bryer T (2007) Toward a Relevant Agenda for a Responsive Public Administration. *JPART* 17(3), 479–500.
Card N (2015) *Applied Meta-Analysis for Social Science Research*. The Guilford Press, New York.
Carroll C, Booth A (2015) Quality Assessment of Qualitative Evidence for Systematic Review and Synthesis: Is It Meaningful, and If So, How Should It Be Performed? *Research Synthesis Methods* 6(2), 149–154.
Christian C (2017) Enhanced Enforcement Outcomes through a Responsive Regulation Approach to Sales Tax Enforcement. *Journal of Public Budgeting, Accounting & Financial Management* 29(4), 464–497.
Coglianese C (2011) Process Choise. *Regulation & Governance* 5(2), 250–261.
Cooke A, Smith D, Booth A (2012) Beyond PICO: The Spider Tool for Qualitative Evidence Synthesis. *Qualitative Health Research* 22(10), 1435–1443.
Cooper H (2017) *Research Synthesis and Meta-Analysis: A Step-by-Step Approach*, 5th edn. SAGE, Los Angeles, CA.
Cooper H, Hedges L, Valentine J (eds) (2019) *The Handbook of Research Synthesis and Meta-Analysis*, 3rd edn. Russel Sage Foundatation, New York.
Curtis A, Miller P, Droste N, McFarlane E, Martino F, Palmer D (2016) "The Ones That Turn Up Are the Ones That Are responsible": Key Stakeholders Perspectives on Liquor Accords. *Drug and Alcohol Review* 35(3), 273–279.
Daniels B (2012) When Agencies Go Nuclear: A Game Theoretic Approach to the Biggest Sticks in an agency's Arsenal. *George Washington Law Review* 80(2), 442–504.
van Duin S, Dekker H, Wielhouwer J, Mendoza J (2018) The Tone from Above: The Effect of Communicating a Supportive Regulatory Strategy on Reporting Quality. *Journal of Accounting Research* 56(2), 467–519.
Eklund Karlson L, Takahashi R (2017) *A Resource for Developing an Evidence Synthesis Report for Policy-Making*. WHO Regional Office for Europe, Copenhagen.
van Erp J (2011) Naming without Shaming: The Publication of Sanctions in the Dutch Financial Market. *Regulation & Governance* 5(3), 287–308.
Eungkyoon L, Jung CS, Kwak J (2016) The Role of Trade Associations in Environmental Compliance under Limited Enforcement: The Case of Small Businesses. *Environmental Policy and Governance* 26(5), 442–436.
Even D (2002) Replication, Meta-Analysis, Scientific Progress, and AMJ's Publication Policy. *Academy of Management Journal* 45(5), 841–846.
Franco A, Malhotra N, Simonovits G (2014) Publication Bias in the Social Sciences: Unlocking the File Drawer. *Science* 345 (6203), 1502–1505.
Freese J, Peterson D (2018) Replication in Social Science. *Annual Review of Sociology* 43(1), 147–165.
Gough D, Oliver S, Thomas J (2012) *An Introduction to Systematic Reviews*. Sage, London.
Gunningham N, Sinclair D (2009) Regulation and the Role of Trust: Reflections from the Mining Industry. *Journal of Law and Society* 36(2), 167–194.
Gurevitch J, Koricheva J, Nakagawa S, Stewart G (2018) Meta-Analysis and the Science of Research Synthesis. *Nature* 555 (7695), 175–182.
Hardwicke TE, Serghiou S, Janiaud P *et al*. (2020) Calibrating the Scientific Ecosystem through Meta-Research. *Annual Review of Statistics and Its Application* 7(1), 1–27.
Hedges J, Bird H, Gilligan G, Godwin A, Ramsay I (2017) The Policy and Practice of Enforcement of Directors' Duties by Statutory Agencies in Australia: An Empirical Analysis. *Melbourne University Law Review* 40(3), 905–966.
Heimer C (2011) Disarticulated Responsiveness: The Theory and Practice of Responsive Regulation in Multi-Layered Systems. *University of British Columbia Law Review* 44(2), 663.
Heyvaert M, Hannes K, Onghena P (2017) *Using Mixed Methods Research Synthesis for Literature Reviews*. SAGE, Los Angeles, CA.
Ho D (2017) Does Peer Review Work? An Experiment of Experimentalism. *Stanford Law Review* 69(1), 1–119.
von Holderstein Holterman J (2009) Outlining the Shadow of the Axe – On Restorative Justice and the Use of Trial and Punishment. *Criminal Law and Philosophy* 3(2), 187–207.
Howell Major C, Savin-Baden M (2010) *An Introduction to Qualitative Research Synthesis: Managing the Information Explosion in Social Science Research*. Routledge, Abingdon.
Hutter B (2001) Is Enforced Self-Regulation a Form of Risk Taking?: The Case of Railway Health and Safety. *International Journal of the Sociology of Law* 29(4), 379–400.
Islam MA, McPhail M (2011) Regulating for Corporate Human Rights Abuses: The Emergence of Corporate Reporting on the ILO's Human Rights Standards within the Global Garment Manufacturing and Retail Industry. *Critical Perspectives on Accounting* 22(8), 790–810.

Ivec M, Braithwaite V, Woods C, Job J (2015) *Applications of Responsive Regulation Theory in Australia and Overseas: Update. Regulatory Institutions Network Occasional Paper 23*. Regulatory Institutions Network, Canberra.
Job J, Stout A, Smith R (2007) Culture Change in Three Taxation Administrations: From Command-and-Control to Responsive Regulation. *Law & Policy* 29(1), 84–101.
Kahneman D (2011) *Thinking Fast and Slow*. Farrar, Straus and Giroux, New York.
Kingsford Smith D (2011) A Harder Nut to Crack? Responsive Regulation in the Financial Services Sector. *University of British Columbia Law Review* 44(3), 695.
Lehmann Nielsen V (2006) Are regulators responsive? *Law & Policy* 28(3), 395–416.
Lehmann Nielsen V, Parker C (2009) Testing Responsive Regulation in Regulatory Enforcement. *Regulation & Governance* 3 (4), 376–399.
Leviner S (2008) An Overview: A New Era of Tax Enforcement – From "Big Stick" to Responsive Regulation. *Regulation & Governance* 2(3), 360–380.
Lewis S, Crawford A, Traynor P (2017) Nipping Crime in the Bud? The Use of Antisocial Behaviour Interventions with Young People in England and Wales. *British Journal of Criminology* 57(5), 1230–1248.
Lochner T, Cain B (1999) Equity and Efficacy in the Enforcement of Campaign Finance Laws. *Texas Law Review* 77(7), 1891–1942.
Lochner T, Apollonio D, Tatum R (2008) Wheat from Chaff: Third-Party Monitoring and Fec Enforcement Actions. *Regulation & Governance* 2(2), 216–233.
Loken E, Gelman A (2017) Measurement Error and the Replication Crisis. *Science* 355(6325), 584–585.
Mahood Q, Van Eerd D, Irvin E (2014) Searching for Grey Literature for Systematic Reviews: Challenges and Benefits. *Research Synthesis Methods* 5(3), 221–234.
Mascini P (2013) Why Was the Enforcement Pyramid So Influential? And What Price Was Paid? *Regulation & Governance* 7 (1), 48–60.
Mascini P, van Wijk E (2009) Responsive Regulation at the Dutch Food and Consumer Product Safety Authority: An Empirical Assessment of Assumptions Underlying the Theory. *Regulation & Governance* 3(1), 27–47.
May P, Winter S (1999) Regulatory Enforcement and Compliance: Examining Danish Agro-Environmental Policy. *Journal of Policy Analysis and Management* 18(4), 625–651.
May P, Wood RS (2003) At the Regulatory Front Lines: Inspectors' Enforcement Styles and Regulatory Compliance. *Journal of Public Administration Research and Theory* 13(2), 117–139.
Mayo-Wilson E, Grant S (2019) Transparent Reporting: Registrations, Protocols, and Final Reports. In: Cooper H, Hedges L, Valentine J (eds) *The Handbook of Research Synthesis and Meta-Analysis*, 3rd edn, pp. 471–488. Russel Sage Foundation, New York.
McKay S, Murray M, MacIntyre S (2015) Pitfalls in Protection: How Theory Can Enrich our Understanding of Regulatory Compliance Problems in Planning Practice. *International Planning Studies* 20(3), 270–291.
Melendez-Torres G, O'Marra-Eves A, Thomas J, Brunton G, Caird J, Petticrew M (2017) Interpretive Analysis of 85 Systematic Reviews Suggests that Narrative Syntheses and Meta-Analyses Are Incommensurate in Argumentation. *Research Synthesis Methods* 8(1), 109–118.
Moss B (2019) Pyramids or Sandcastles? Regulating Police Conduct in Ireland. *Policing and Society* 20(2), 1027–1043.
Muehlbacher S, Kirchler E, Schwarzenberger H (2011) Voluntary Versus Enforced Tax Compliance: Empirical Evidence for the "Slippery Slope" Framework. *European Journal of Law and Economics* 32(1), 89–97.
National Academies of Sciences, Engineering, and Medicine (2019) *Reproducibility and Replicability in Science*. The National Academies Press, Washington, DC.
Nonet P, Selznick P (2009 [1978]) *Law & Society in Transition: Toward Responsive Law*. Transaction Publishers, New Brunswick.
O'Neill B (2018) Corporate Crime and Regulatory Discretion: Rethinking the Use of Criminal, Civil and Administrative Penalties. *Criminal Law Journal* 42(5), 322–336.
Parker C (2006) The "Compliance" Trap: The Moral Message in Responsive Regulatory Enforcement. *Law & Society Review* 40(3), 591–622.
Parker C (2013) Twenty Years of Responsive Regulation: An Appreciation and Appraisal. *Regulation & Governance* 7(1), 1–151.
Pawson R (2002) Evidence-Based Policy: The Promise of "Realist Synthesis". *Evaluation* 8(3), 340–358.
Petticrew M, Roberts H (2006) *Systematic Reviews in the Social Sciences: A Practical Guide*. Blackwell, Oxford.
Pires R (2008) Promoting Sustainable Compliance: Styles of Labour Inspection and Compliance Outcomes in Brazil. *International Labour Review* 147(2–3), 199–229.
Quilter J, McNamara L (2015) "Long May the Buskers Carry on busking": Street Music and the Law in Melbourne and Sydney. *Melbourne University Law Review* 39(2), 539–591.
Ringquist E (2013) *Meta-Analysis for Public Management and Policy*. John Wiley, San Francisco, CA.
Saini M, Shlonsky A (2012) *Systematic Synthesis of Qualitative Research*. Oxford University Press, Oxford.
Sampson H, Bloor M (2007) When Jack Gets out of the Box: The Problems of Regulating a Global Industry. *Sociology* 41(3), 551–569.
Schooler J (2014) Metascience Could Rescue the "Replication Crisis". *Nature* 515(7525), 9.
Shaddish W, Lecy J (2015) The Meta-Analytic Big Bang. *Research Synthesis Methods* 6(3), 246–264.
Sheehy B, Feaver D (2015) Designing Effective Regulation: A Normative Theory. *UNSW Law Journal* 38(1), 392–425.
Short JL (2021) The politics of regulatory enforcement and compliance: Theorizing and operationalizing political influences. *Regulation & Governance* 15(3), 653–685. http://dx.doi.org/10.1111/rego.12291.

Short J, Toffel MW (2010) Making Self-Regulation More Than Merely Symbolic: The Critical Role of the Legal Environment. *Administrative Science Quarterly* 55(2), 361–396.
Silbey S (1984) The Consequences of Responsive Regulation. In: Hawkins K, Thomas JM (eds) *Enforcing Regulation*, pp. 147–170. Kluwer Nijhodd, Boston, MA.
Simpson S (2013) White-Collar Crime: A Review of Recent Developments and Promising Directions for Future Research. *Annual Review of Sociology* 39(1), 309–331.
Smith S (2015) *Taxation: A Very Short Introduction*. Oxford University Press, Oxford.
Stoll C, Izadi S, Fowler S, Green P, Suts J, Coldits G (2019) The Value of a Second Reviewer for Study Selection in Systematic Reviews. *Research Synthesis Methods* 10(4), 539–545.
Sutton A, Papaioannou D, Booth A (2016) *Systematic Approaches to a Successful Literature Review*. Sage, London.
Szejnwald Brown H, Angel D, Broszkiewicz R, Krzyśków B (2001) Occupational Safety and Health in Poland in the 1990s: A Regulatory System Adapting to Societal Transformation. *Policy Sciences* 34(3/4), 247–271.
Toffel MW, Short J (2011) Coming Clean and Cleaning up: Does Voluntary Self-Reporting Indicate Effective Self-Policing? *Journal of Law and Economics* 54(4), 609–649.
Tombs S, Whyte D (2013) Transcending the Deregulation Debate? Regulation, Risk, and the Enforcement of Health and Safety Law in the Uk. *Regulation & Governance* 7(1), 61–79.
Tränkle S (2007) In the Shadow of Penal Law: Victim–Offender Mediation in Germany and France. *Punishment & Society* 9 (4), 395–415.
Valentine J (2019) Incorporating Judgments about Study Quality into Research Syntheses. In: Cooper H, Hedges L, Valentine J (eds) *The Handbook of Research Synthesis and Meta-Analysis*, 3rd edn, pp. 129–140. Russel Sage Foundation, New York.
Van der Heijden J (2010) A Short History of Studying Incremental Institutional Change: Does *Explaining Institutional Change* Provide any New Explanations? *Regulation & Governance* 4(2), 230–243.
Vevea J, Coburn K, Sutton A (2019) Publication Bias. In: Cooper H, Hedges L, Valentine J (eds) *The Handbook of Research Synthesis and Meta-Analysis*, 3rd edn, pp. 383–429. Russel Sage Foundation, New York.
Walker-Munro B (2019) Regulating Disruption and Development of the Disruption Calculus. *University of Western Australia Law Review* 46(1), 111–143.
Waller V (2007) The Challenge of Institutional Integrity in Responsive Regulation: Field Inspections by the Australian Taxation Office. *Law & Policy* 29(1), 67–83.
Welsh M (2009) Civil Penalties and Responsive Regulation: The Gap between Theory and Practice. *Melbourne University Law Review* 33(3), 908–933.
Wilson E (1999) *Consilience: The Unity of Knowledge*. Vintage Books, New York.
Wilson D (2019) Systematic Coding for Research Synthesis. In: Cooper H, Hedges L, Valentine J (eds) *The Handbook of Research Synthesis and Meta-Analysis*, 3rd edn, pp. 153–172. Russel Sage Foundation, New York.
Winter S, May P (2001) Motivation for Compliance with Environmental Regulations. *Journal of Policy Analysis and Management* 20(4), 675–698.
Woods C, Ivec M, Job J, Braithwaite V (2010) *Applications of Responsive Regulation Theory in Australia and Overseas. Regulatory Institutions Network Occasional Paper 15*. Regulatory Institutions Network, Canberra.
Yee W-H, Tang S-Y, Lo CW-H (2016) Regulatory Compliance When the Rule of Law Is Weak: Evidence from China's Environmental Reform. *Journal of Public Administration Research and Theory* 26(1), 95–112.
Zhu J, Chertow MR (2019) Authoritarian but Responsive: Local Regulation of Industrial Energy Efficiency in Jiangsu, China. *Regulation & Governance* 13(3), 384–404.

Supporting information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Appendix S1. A narrative review of selected meta-research publications
Appendix S2. Meta-research protocol
Appendix S3. Responsive regulation terminology for the inclusion and exclusion process
Appendix S4. Flowchart of the sourcing, inclusion, and exclusion of studies
Appendix S5. Quality assessment
Appendix S6. Codes used for analyzing included publications

