



# Kent Academic Repository

**Shan, Yafeng and Williamson, Jon (2021) *Applying Evidential Pluralism to the social sciences*. European Journal for Philosophy of Science, 11 (4). pp. 1-27. ISSN 1879-4912.**

## Downloaded from

<https://kar.kent.ac.uk/89911/> The University of Kent's Academic Repository KAR

## The version of record is available from

<https://doi.org/10.1007/s13194-021-00415-z>

## This document version

Publisher pdf

## DOI for this version

## Licence for this version

UNSPECIFIED

## Additional information

## Versions of research works

### Versions of Record

If this version is the version of record, it is the same as the published version available on the publisher's web site. Cite as the published version.

### Author Accepted Manuscripts

If this document is identified as the Author Accepted Manuscript it is the version after peer review but before type setting, copy editing or publisher branding. Cite as Surname, Initial. (Year) 'Title of article'. To be published in *Title of Journal*, Volume and issue numbers [peer-reviewed accepted version]. Available at: DOI or URL (Accessed: date).

## Enquiries

If you have questions about this document contact [ResearchSupport@kent.ac.uk](mailto:ResearchSupport@kent.ac.uk). Please include the URL of the record in KAR. If you believe that your, or a third party's rights have been compromised through this document please see our [Take Down policy](https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies) (available from <https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies>).



# Applying Evidential Pluralism to the social sciences

Yafeng Shan<sup>1</sup> · Jon Williamson<sup>1</sup>

Received: 6 November 2020 / Accepted: 9 September 2021 / Published online: 08 October 2021  
© The Author(s) 2021

## Abstract

Evidential Pluralism maintains that in order to establish a causal claim one normally needs to establish the existence of an appropriate conditional correlation and the existence of an appropriate mechanism complex, so when assessing a causal claim one ought to consider both association studies and mechanistic studies. Hitherto, Evidential Pluralism has been applied to medicine, leading to the EBM+ programme, which recommends that evidence-based medicine should systematically evaluate mechanistic studies alongside clinical studies. This paper argues that Evidential Pluralism can also be fruitfully applied to the social sciences. In particular, Evidential Pluralism provides (i) a new approach to evidence-based policy; (ii) a new account of the evidential relationships in more theoretical research; and (iii) new philosophical motivation for mixed methods research. The application of Evidential Pluralism to the social sciences is also defended against two objections.

**Keywords** Evidential Pluralism · Russo-Williamson thesis · Evidence-based policy · Mixed methods research · Social science methodology · Causality · Causation

This paper makes the case that a specific account of the epistemology of causality, namely Evidential Pluralism, can be fruitfully applied to the social sciences. Section 1 provides a brief introduction to Evidential Pluralism and its application to medicine. Section 2 argues that there is scope to apply Evidential Pluralism to the social sciences, in particular, to evidence-based policy, to theoretical social science research, and to mixed methods research. We then respond to two objections to the claim that

---

✉ Yafeng Shan  
y.shan@kent.ac.uk

Jon Williamson  
j.williamson@kent.ac.uk

<sup>1</sup> Department of Philosophy and Centre for Reasoning, University of Kent, Canterbury, UK

Evidential Pluralism can be applied within the social sciences: one due to Julian Reiss (Section 3) and a second due to Francois Claveau (Section 4). In Section 4 we also see how Evidential Pluralism applies to single-case causal claims in the social sciences. We conclude in Section 5 that Evidential Pluralism sheds new light on the use of evidence in the social sciences.

## 1 What is Evidential Pluralism?

That correlation does not imply causation was already well understood by the end of the 19th century. One influential response to this platitude was to replace talk of causation by talk of correlation. Thus Yule (1895, p. 605) says, ‘We have found that the rate of total pauperism is *positively* correlated with the proportion of out-relief given . . . This statement does not say either that the low mean proportion of the out-relief is the *cause* of the lesser pauperism or *vice versa*: such terms seem best avoided.’ These days, however, this response is recognised to be unsatisfactory, because it is usually causation that we are interested in, not correlation. It was important to determine whether the proportion of out-relief (aid to the poor outside a workhouse) was a cause of the total level of pauperism, for only if the relationship were causal would a change to welfare policy be warranted, and only if the relationship were causal could the proportion of out-relief explain the level of pauperism.

That correlation does not imply causation is witnessed by correlations such as those between British bread prices and the sea level in Venice (Sober, 2001) or between the supply of shrimp in the US and the number of people killed by sharp glass (Vigen, 2015). As is apparent from Table 1, there are a large number of possible explanations of an observed correlation between variables *A* and *B*, only one of which is that *A* is a cause of *B*. In order to establish causation, one must rule out these other possible explanations.

Fortunately, there is something distinctive about those correlations that are genuinely causal. If it is genuinely the case that an observed correlation is attributable to the fact that *A* is a cause of *B*, then there is some combination of mechanisms which explains instances of *B* by invoking instances of *A* and which can account for the magnitude of the observed correlation. It is the existence of this mechanism from *A* to *B* that distinguishes causation from other potential explanations of a correlation. Thus Russo and Williamson (2007) put forward the thesis that, in order to establish a causal claim, one normally needs to establish both the existence of an appropriate correlation and the existence of mechanism complex that can account for the correlation. While Russo and Williamson restricted this thesis to medicine and the health sciences, in their closing remarks they suggested that their thesis might also be applicable beyond medicine, to the social sciences (Russo & Williamson, 2007, p. 169). In this paper we follow through on this suggestion, by extending the thesis to the social sciences. We will not attempt a general defence of this thesis here, nor of its application to medicine, as this has been discussed at length elsewhere—see, e.g., Williamson (2019a) and references therein. Instead we shall explain how this thesis motivates Evidential Pluralism and, in subsequent sections, how Evidential Pluralism can be fruitfully applied to the social sciences.

**Table 1** Possible explanations of an observed correlation between *A* and *B*

Causation	<i>A</i> is a cause of <i>B</i> .
Reverse causation	<i>B</i> is a cause of <i>A</i> .
Confounding (selection bias)	There is some confounder <i>C</i> that has not been adequately controlled for by the study.
Performance bias	Those in the <i>A</i> -group are identified and treated differently to those in the $\neg A$ -group.
Detection bias	<i>B</i> is measured differently in the <i>A</i> -group in comparison to the $\neg A$ -group.
Chance	Sheer coincidence, attributable to too small a sample.
Fishing	Measuring so many outcomes that there is likely to be a chance correlation between <i>A</i> and some such <i>B</i> .
Temporal trends	<i>A</i> and <i>B</i> both increase over time for independent reasons. E.g., prevalence of coeliac disease & spread of HIV.
Semantic relationships	Overlapping meaning. E.g., phthisis, consumption, scrofula (all of which refer to tuberculosis).
Constitutive relationships	One variable is a part or component of the other.
Logical relationships	Measurable variables <i>A</i> and <i>B</i> are logically complex and logically overlapping. E.g., <i>A</i> is $C \wedge D$ and <i>B</i> is $D \vee E$ .
Physical laws	E.g., conservation of total energy can induce a correlation between two energy measurements
Mathematical relationships	E.g., mean and variance variables from the same distribution will often be correlated.

From Williamson (2019a)

The thesis of Russo and Williamson (2007) concerns what one needs to establish in order to establish causation: the existence of a correlation and the existence of a mechanism. One can also distinguish the kinds of study used to establish correlation and mechanism (Illari, 2011; Williamson, 2021c, Section 3). The obvious way to establish a correlation is to perform a statistical study, which samples a large number of individuals, measuring *A* and *B* and assessing the extent to which they occur together. In medicine and the health sciences, these studies include clinical and epidemiological studies and they often measure *A* and *B* together with some set *C* of potential confounders, in order to determine whether *A* and *B* are probabilistically dependent conditional on *C*.<sup>1</sup> We shall refer to any study that tests for an association between *A* and *B* by measuring *A* and *B* together as an *association study*. This class of studies includes both experimental studies and observational studies.

On the other hand, the obvious way to determine whether there is a suitable mechanism complex by which instances of *A* are responsible for instances of *B* is to carry out studies that look for key features of the putative mechanisms linking *A* and *B*, including intermediary variables as well as the entities and activities that constitute the mechanism, and the way these constituents are organised. We shall refer to such

<sup>1</sup>A potential confounder is a variable that, so far as one can tell from the available evidence, plausibly might account for the correlation between *A* and *B*—e.g., a common cause of *A* and *B*.

a study as a *mechanistic study*. While it is in principle possible to design a study that is both an association study and a mechanistic study, mechanistic studies typically do not measure *A* and *B* together, but instead seek to shed light on component links or parts of the mechanism complex. Thus there is a sense in which mechanistic studies typically offer a fine-grained view of the putative causal relationship, while association studies typically offer a coarse-grained perspective.<sup>2</sup>

These evidential relationships are captured by Fig. 1. Association studies can be used to directly test the hypothesis that the putative cause and effect are correlated conditional on potential confounders ( $C_1$ ). In certain circumstances, they can also indirectly confirm the claim that there is a mechanism responsible for this correlation ( $C_2$ ): for example, if all the association studies find a large correlation and all potential confounders have been adequately controlled for, and other explanations of this correlation (see Table 1) can be ruled out, then one might infer that there must be some mechanism of action that gives rise to the correlation. Alternatively, one can confirm the existence of an appropriate mechanism complex by means of specific mechanism hypotheses which posit features of the various mechanisms linking *A* and *B* ( $M_2$ ). The presence of these features can then be tested by mechanistic studies ( $M_1$ ). In certain circumstances, specific mechanism hypotheses can also support the claim that the variables of interest are genuinely correlated ( $M_3$ ): for example, the features of a parachute mechanism confirm the claim that, when falling from high altitude, parachute use is negatively correlated with serious injury, obviating the need for randomised trials as evidence of causation (Williamson, 2019a, §2.2).

Figure 1 encapsulates the content of Evidential Pluralism:

**Evidential Pluralism.** In order to establish a causal claim one normally needs to establish the existence of an appropriate conditional correlation and the existence of an appropriate mechanism complex, so when assessing a causal claim one ought to consider relevant association studies and mechanistic studies, where available.

Some points of clarification. Firstly, Evidential Pluralism is more specific than the claim that one should consider all relevant evidence when assessing the truth of a proposition (aka the Principle of Total Evidence). This is because Evidential Pluralism specifies which kinds of study are relevant and why they are relevant. Similarly, Evidential Pluralism is more specific than the claim that it is desirable to use diverse methods, theories or perspectives in research (aka ‘triangulation’): Evidential Pluralism holds that a very particular kind of triangulation is required for causal inference.

Second, Evidential Pluralism is a purely epistemological thesis: it is a thesis about establishing and assessing causality, not an analysis of the concept of cause nor a

<sup>2</sup>It should be emphasised that a mechanism complex linking *A* and *B* may be composed of multiple mechanisms, some of which may counteract others. For example, there are multiple mechanisms linking the legalisation of abortion to the crime rate, as we shall see in Section 2.2.

Note too, that the distinction between an association study and a mechanistic study is defined relative to the variables of interest *A* and *B*. A study that finds some association between *A* and *D*, where *D* is a purported intermediary variable on a mechanism from *A* to *B* would count as a mechanistic study relative to *A* and *B* but an association study relative to *A* and *D*.

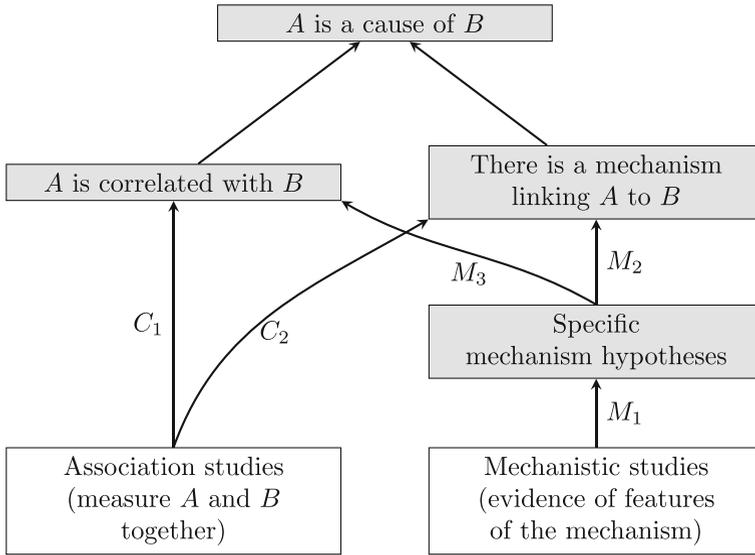


Fig. 1 Evidential relationships for establishing a causal claim (Williamson, 2021c)

claim about the metaphysical nature of causality.<sup>3</sup> Similarly, Evidential Pluralism makes no direct claims about the nature of mechanisms, although it does appeal to the concept of mechanism. From the point of view of the epistemology of causality, it is important to distinguish evidence of mechanisms from evidence of correlation regardless of whether mechanisms themselves are ultimately reducible to correlations, laws, dispositions, low-level causal relations, or none of these. Marchionni and Reijula (2019), for example, suggest that mechanisms are reducible to chains of difference making involving intermediary variables, while Ioannidis and Psillos (2017) analyse mechanisms as causal pathways—a view that they argue requires minimal metaphysical commitments.<sup>4</sup> Such accounts are fully compatible with the two distinctions to which Evidential Pluralism appeals: the distinction between a correlation between *A* and *B* and a mechanism complex linking *A* and *B*, and the distinction between an association study and a mechanistic study. Only by drawing these

<sup>3</sup>To say that Evidential Pluralism is a purely epistemological thesis does not imply that it is entirely devoid of metaphysical consequences. As Russo and Williamson (2007) and Weber (2009) observe, some metaphysical theories of causality may better accommodate Evidential Pluralism than others, in which case Evidential Pluralism favours the former theories over the latter.

<sup>4</sup>Hedström and Ylikoski (2010, p. 51) criticise the view that mechanisms in the social sciences can be understood solely in terms of intermediary variables. While we make no attempt to settle the question of how best to analyse mechanisms here, we should note that the literature on Evidential Pluralism often takes a mechanism complex to be composed of complex systems mechanisms and/or mechanistic processes. A complex-systems mechanism consists of entities and activities organised in such a way that they are responsible for some phenomenon (Machamer et al., 2000; Illari & Williamson, 2012). A mechanistic process is a spatiotemporally contiguous process along which a signal can be propagated (Reichenbach, 1956; Salmon, 1998).

distinctions can one properly understand the evidential relationships involved in the assessment of causality.

Third, Evidential Pluralism is primarily a normative thesis, about how one ought to assess a causal claim. As to whether causal claims are actually assessed on this basis varies. The methods of the International Agency for Research on Cancer, which is charged with assessing carcinogenicity claims, are arguably compatible with Evidential Pluralism (Williamson, 2019b), while certain implementations of present-day evidence-based medicine (EBM) are not (Williamson, 2019a). Present-day EBM presupposes a kind of evidential monism: it focusses almost exclusively on association studies—especially randomised controlled clinical trials (RCTs)—and tends to either ignore mechanistic studies or to view mechanistic studies as inherently low-quality evidence, to be trumped by association studies. Parkkinen et al. (2018) argue against the monistic approach of present-day EBM. They develop the EBM+ programme, which appeals to Evidential Pluralism as grounds for systematically assessing mechanistic studies alongside association studies, thereby extending the evidence base that is scrutinised by EBM. Indeed, the main novelty of Evidential Pluralism is its dual focus on association studies and mechanistic studies: many other approaches to causal evaluation presuppose the primacy of one or other kind of evidence, or are more radically pluralist.

Fourth, in the above formulation of Evidential Pluralism, the qualification ‘normally’ is used because in the philosophical literature there are well-known examples of causation without correlation and of causation without mechanism see (Williamson, 2019a, § 1.1.5). A detailed consideration of these pathological cases would take us too far from the main argument of this paper, so we simply set them aside here.

Fifth, when we say that one ought to ‘consider’ a study we mean *explicitly and systematically assess* the study, unless the study only bears only on propositions that have already been established, in which case one only need to take those propositions into account, not the studies that helped to establish those propositions. Thus one usually only needs to scrutinise studies that have a bearing on what is not already established. There is, however, an exception to this rule: if new studies overturn previously established claims, then the studies that support those claims may need to be revisited.

Finally, it is important to be clear what it is to establish the existence of a mechanism. Establishing a proposition enables it to be used as evidence for other propositions, and thus requires a substantial body of evidence (Williamson, 2021a). To establish the existence of a mechanism it is not enough merely to have a story about what the mechanism might be. As mentioned above, one way to establish the existence of a mechanism is by confirming the key features of the hypothesised mechanism (channels  $M_1$  and  $M_2$  in Fig. 1), while another proceeds indirectly, by ruling out alternative explanations of an observed correlation (channel  $C_2$ ). If the former route is taken, it is not enough to establish the existence of *some* pathway of action: it is also essential to establish that the influence of this pathway is not negated by that of counteracting pathways. Thus one needs to consider the mechanism complex as a whole that links  $A$  and  $B$ . Moreover, it is not enough to show the purported

mechanism merely exists in some individuals—it needs to be present in enough individuals to be able to account for the extent of the observed correlation.

## 2 Applying Evidential Pluralism to the social sciences

In this section we argue that there are several ways in which Evidential Pluralism can be fruitfully applied to the social sciences. Most obviously, the move from EBM to EBM+ warrants an analogous move from present-day evidence-based policy (EBP) to EBP+, a new approach to policy appraisal which takes evidence of mechanisms more seriously (Section 2.1). Of course, causal claims in the social sciences are not limited to claims about the effectiveness of proposed policy interventions—they also include claims about the causes and effects of societal, economic, legal, geographical, linguistic and psychological phenomena, for example. These claims are at the heart of what might be called ‘basic social science research’. We argue that Evidential Pluralism can be usefully applied to basic social science research, in addition to policy appraisal, because it sheds new light on the evidential relationships involved in establishing causation (Section 2.2). The question then arises as to how Evidential Pluralism relates to mixed methods research in the social sciences. We argue in Section 2.3 that there are important differences between the conceptual distinctions underlying Evidential Pluralism and those prevalent in mixed methods research, but that Evidential Pluralism can nevertheless be thought of as providing new, metaphysics-free, philosophical foundations for certain kinds of mixed methods research.

### 2.1 From evidence-based medicine to evidence-based policy

In the 1990s, the methods of EBM quickly spread to the evaluation of social interventions, leading to what is now known as evidence-based policy (EBP). The Cochrane Collaboration, which promotes EBM, was set up in 1993, while the Campbell Collaboration, which promotes EBP along similar lines, was created in 1999.

In the UK, for example, the primary organ of EBP is the government-led ‘What Works Network’, which includes the National Institute of Health and Care Excellence (NICE) as well as eight other centres tasked with evaluating social interventions. The What Works Network is built around the use of association studies (in particular, RCTs) as the evidence on which to base an evaluation (What Works, 2018a, p. 4) and membership of the network is restricted to centres which share the ranking of evidence promulgated by present-day EBM/EBP (Cabinet Office, 2018, p. 3). Allied to the What Works Network is the UK Government Trials Advice Panel, which was set up in 2015 to promote the use of RCTs in public policy decision making (What Works, 2018b). These structures ensure that the monistic methods of present-day EBM/EBP are entrenched at the heart of policy making in the UK.

The situation in the UK is just one instance of a global phenomenon. In the US, the dominant approach to EBP is also modelled on EBM, with a focus on RCTs (Baron, 2018). In addition, the United Nations actively promotes a global vision of EBP based on statistical association studies—see, e.g., United Nations (2013). EBP and EBM

continue to develop hand-in-hand: e.g., the Cochrane and Campbell Collaborations share methods at events such as the ‘Global Evidence Summit’, held in 2017 in Cape Town and 2023 in Prague.

Thus, EBP is modelled on EBM, which, as we noted in Section 1, underestimates the importance of mechanistic studies. Given this, there is arguably a need for what we will call ‘EBP+’, i.e., an analogue of EBM+, but applied to policy evaluation. As with EBM+, the aim of EBP+ is to provide methods for systematically assessing mechanistic studies and integrating these assessments with those of association studies in order to determine the status of a causal claim. The need for EBP+ arises because Evidential Pluralism applies equally to medicine and policy making. If Fig. 1 captures the key evidential relationships when evaluating a causal claim in medicine, then it does so too in the social sciences, in particular when evaluating a social intervention. Moreover, EBP+ is also required for successful extrapolation: in both medicine and the social sciences, it is only by considering mechanisms of action that one can decide whether a causal relationship discovered in a study context can be extrapolated to a target context of application (see, e.g., Steel, 2008; Wilde & Parkkinen, 2019).<sup>5</sup> Thus EBP+ would proceed very much along the same lines as EBM+. The evaluation methods common to both are set out in detail by Parkkinen et al. (2018).

One might wonder whether there is some systematic difference between the social sciences and the biomedical sciences that undermines the applicability of Evidential Pluralism to the social sciences. However, one cannot draw a sharp distinction between the evaluation of causal claims in medicine and those in the social sciences. Indeed, such claims often overlap: health policy interventions are interventions in both medicine and social policy; hence the inclusion of NICE in the What Works Network. Insofar as one can generalise, the main methodological difference between the biomedical sciences and the social sciences is that in the social sciences it can be harder to isolate an experiment from contextual factors that might influence its results and that can thwart replication. In many cases it can also be harder to properly randomise individuals to social policy interventions, to construct a placebo intervention for the control group, and to ensure adherence to the social policy interventions being tested. In addition, it can be harder to successfully extrapolate claim about the effectiveness of a social intervention from one setting to another than to extrapolate a claim about the effectiveness of a medical intervention. This is because social settings can vary much more widely in their social mechanisms than do human bodies in their pathophysiological mechanisms. All these considerations favour a shift towards Evidential Pluralism over the current reliance on RCTs.

Some EBP practitioners have indeed begun to question the current focus on RCTs. For example, Yamey and Feachem (2011) observe that,

---

<sup>5</sup>Cartwright and Hardie (2012, Section A.1.1) provide an nice example of the importance of extrapolation to public health policy. They also criticise the way in which EBP focusses almost exclusively on association studies—in particular on RCTs. Evidential Pluralism can be thought of as a principled way of addressing some of these concerns with extrapolation and RCTs. Parkkinen et al. (2018) provide a detailed account of the logic of extrapolation that is motivated by Evidential Pluralism.

while the RCT is rightly hailed as the ‘pinnacle’ of evidence-based medicine, in the global public health community, there is growing recognition that new research designs are desperately needed to help evaluate ‘real world’ programmes. Such designs would, we believe, also help to illuminate the implementation ‘black box’. (Yamey & Feachem, 2011, p. 98.)

Moreover, one of the What Works centres has begun to recognise the importance of mechanisms. The What Works Centre for Crime Reduction has developed the ‘EMMIE’ framework for systematic reviews of evidence: Effect size, Mechanism, Moderator, Implementation and Economics are all components of an evaluation (Johnson et al., 2015; Tilley, 2016; Thornton et al., 2019). From the point of view of Evidential Pluralism, considering mechanisms is an important step in the right direction. However, EMMIE is based not on Evidential Pluralism but on the realist evaluation approach of Pawson and Tilley (1997), and it will be instructive to consider how their approach differs from one based on Evidential Pluralism.

While mechanisms are important to both realist evaluation and Evidential Pluralism, there are three key philosophical differences between these two approaches.

First, the realist evaluation approach of Pawson and Tilley (1997) makes a firm commitment to scientific realism: specifically, a causal powers metaphysics of causation (Pawson & Tilley, 1997, pp. 33, 56). This realism is accompanied by a rejection of Humean and Kantian metaphysics, which hold that causal relationships are a device we employ to structure the world, and which do not posit causal powers or causal necessitation ‘out there’ in the world. Evidential Pluralism, in contrast, is a purely epistemological thesis that makes no specific metaphysical claims. It is compatible with an anti-realist account of causation which analyses causal claims in terms of rational beliefs, for example (Williamson, 2021b, Appendix).

The second philosophical difference between the realist evaluation of Pawson and Tilley (1997) and Evidential Pluralism is that, on account of its metaphysical commitment, their approach involves a rejection of the experimental methodology that underpins RCTs and certain other kinds of association study. Evidential Pluralism, in contrast, takes experimental methods to have the potential to provide good evidence, relevant to the assessment of a causal claim. If Evidential Pluralism is right, one shouldn’t reject these methods—rather, one should augment them, by considering mechanistic studies alongside association studies. The aim of EBP+ would be to improve, rather than overturn, present-day EBP.

Third, realist evaluation proceeds from the premise that there is no logic of evaluation (Pawson & Tilley, 1997, xiii) while Evidential Pluralism takes there to be a logic of evaluation, portrayed by Fig. 1. According to this logic of evaluation, causation is established by establishing correlation and mechanism, which in turn requires assessment of the confirmation channels  $C_1$ ,  $C_2$ ,  $M_1$ ,  $M_2$ ,  $M_3$ , i.e., the assessment of any relevant association studies and mechanistic studies. Parkkinen et al. (2018) show in the context of medicine that this logic of evaluation can be broken down into a series of practical steps.

We should note that realist evaluation has been developed in a number of different directions since 1997—see Jagosh et al. (2016), for example, for some pointers. In particular, not all proponents of realist evaluation now reject the experimental

methodology and RCTs. For example, Bonell et al. (2012) argue for the use of RCTs in a way that is sensitive to the concerns of realist evaluation. Moreover, the development of EMMIE can be considered to be a move towards a logic of evaluation.

Thornton et al. (2019), although proponents of EMMIE, identify some limitations of realist evaluation as implemented in the EMMIE approach. In practice, EMMIE exclusively scrutinises systematic reviews, which almost always consider association studies rather than mechanistic studies, so mechanistic evidence tends to appear rather scant and hence to be rated as weak in EMMIE evaluations. From the point of view of Evidential Pluralism, it is not enough to consider systematic reviews of association studies—it is essential to articulate specific mechanism hypotheses and to search the literature for evidence relevant to those hypotheses. The International Agency for Research on Cancer provides an example of good practice here (Williamson, 2021d): each carcinogenicity evaluation has a dedicated subgroup responsible for systematically assessing mechanistic studies.

Furthermore, since an EMMIE evaluation has five components, it is not obvious how these five aspects should combine to give an overall assessment. This opens the door to subjective judgements of relative importance to influence the overall assessment. Thus, Thornton et al. (2019) worry that realist reviews may not be replicable. This is less of a concern for Evidential Pluralism, which only has two strands to integrate, namely evidence of correlation and evidence of mechanisms (Williamson, 2021d). Parkkinen et al. (2018, Chapter 7) offer a systematic way of integrating these two strands in order to come to an overall assessment.

One final point which is worth noting with respect to EMMIE is that ‘Mechanism’ is graded on a scale from 0 to 4 but only grade 4 requires concrete evidence of mechanism: grades 1 to 3 merely require some story or theory about what the mechanism might be (Thornton et al., 2019, Figure 2). This may stem from the important role of theory in realist evaluation (Pawson & Tilley, 1997, p. 59). In contrast, Evidential Pluralism is concerned with evidence, not theory. In the social sciences, it is often very easy to conjecture a mechanism hypothesis, and a story of a mechanism that is not backed up by evidence has no confirmatory value, for the Evidential Pluralist.

Although Evidential Pluralism differs from realist evaluation in important ways, the two approaches do share some key claims: most notably, that EBP needs to move beyond EBM’s monistic focus on association studies, and that mechanisms should play a prominent role. Evidential Pluralism can be thought of providing some motivation for these claims that is not tied to realism.

There are other methodologies that fit naturally with Evidential Pluralism. For example, process tracing (Mahoney, 2012), contribution analysis (Mayne, 2001), and sociomarkers (Ghiara & Russo, 2019) have all been put forward as ways of developing and evaluating specific mechanism hypotheses and might thus be invoked to help assess channel  $M_1$  in Fig. 1. Moreover, as we shall see shortly, Evidential Pluralism coheres well with a form of mixed methods research. First, though, we turn to the use of Evidential Pluralism in areas of the social sciences other than policy evaluation.

## 2.2 Evidential Pluralism in basic social science research

In medicine, claims about the effects of interventions are not the only causal claims of interest. Claims about the effects of pathogens, chemical exposures and lifestyle factors, for instance, are also central to medicine. More generally, basic medical research is concerned largely with claims about the causal components of mechanisms for health or disease. Likewise, the interests of the social sciences extend well beyond claims about the effects of interventions. The causal components of the mechanisms relevant to the various social sciences are the bread and butter of basic social science research. Figure 1 and Evidential Pluralism apply equally to these causal claims as to claims about interventions. While we do not suggest that social scientists conceptualise their research methodology in terms of Evidential Pluralism, good social science research tends to take both association studies and mechanistic studies into account, where available. (For some examples of mechanistic studies in basic social science research, of varying study designs, see Table 2.) Framing basic social science research in terms of Evidential Pluralism can help researchers to understand the core confirmatory relationships between items of evidence and to reach an overall assessment of the credibility of the causal claim of interest.<sup>6</sup>

Of course, Evidential Pluralism is not the only approach to emphasise the importance of mechanisms to the social sciences. We have already seen that mechanisms are central to realist evaluation, and a mechanism-based approach has also been developed and defended as a central part of analytic sociology (Hedström & Swedberg, 1998; Demeulenaere, 2011). However, Evidential Pluralism differs from this latter mechanism-based approach, just as it differs from realist evaluation. Firstly, the two approaches address different questions: the mechanism-based approach emphasises the role of mechanism-based explanations in the social sciences, whereas Evidential Pluralism concerns the types of evidence needed to establish a causal claim. For this reason, the mechanism-based approach focusses primarily on mechanisms, while Evidential Pluralism is a dualist approach, treating evidence of correlation and evidence of mechanisms on a par. Second, as in the case of realist evaluation, the mechanism-based approach makes metaphysical commitments that are not made by Evidential Pluralism. The mechanism-based approach typically assumes that causal relationships can be analysed or characterised in terms of mechanisms, while according to Evidential Pluralism, mechanisms are merely an important indicator of causality. Moreover, as Hedström and Ylikoski (2010, p. 64) argue, ‘underlying the mechanism-based approach is a commitment to realism and an opposition to any form of instrumentalism.’ Evidential Pluralism makes no explicit metaphysical claims.

As an example of the compatibility of basic social science research with Evidential Pluralism, consider Donohue and Levitt’s study on legalised abortion and crime rates. Donohue and Levitt (2001) argue that the legalisation of abortion in the 1970s was a cause of the decline in the crime rates in the 1990s in the United States. In

---

<sup>6</sup>Additionally, Moneta and Russo (2014) argue that in the context of econometrics, Evidential Pluralism can help to elucidate the difference between statistical models, which merely describe associations, and causal models.

**Table 2** Examples of mechanistic studies in the social sciences

Specific mechanism hypothesis	Study (and study design)
More stringent search requirements (as an intermediary variable between strictness of unemployment benefits and unemployment rates) lower the chance of being re-employed by the same employer.	The 'work search' experiment in the research of Johnson and Klepinger (1994) on the effects of monitoring and sanctions on unemployment. (RCT)
Legalised abortion has a disproportionate effect on the birth of those who are most at risk of engaging in criminal behaviour, which in turn decreases crime rates.	The longitudinal analysis in the research of Donohue and Levitt (2001) on legalised abortion and crime rates. See below for more detail. (Longitudinal study)
The initial endowments to which rebel leaders have access constrain their tactics of recruitment, which shape the membership profile of a rebel group. The membership profile affects its internal organisation and eventually the strategies of violence in war.	The ethnographic interviews in the research of Weinstein (2007) on causes of the patterns of violence in rebels. See below for more detail. (Interview)
Family socioeconomic status and mental abilities affect significant others' influences (e.g., parental influence, teachers' influence, and friends' influence) on a youth and her own observations of her ability, which in turn affect her levels of educational and occupational aspiration, which ultimately influence subsequent levels of educational attainment.	The cross-sectional study in the research of Sewell et al. (1969) on family background and educational attainment. (Cross-sectional study)
Economic interdependence between states influences the expectation that a state has of future trade with other states, which eventually influences the chances of military conflict.	The large-N qualitative study in the research of Copeland (2015) on economic interdependence and military conflict between states. (Large-N qualitative analysis)

order to establish this causal claim, they provide evidence of mechanisms as well as evidence of correlation. They find two mechanisms of action. The first is that legalising abortion reduces crime through smaller cohort sizes. The smaller cohort that results from legalised abortion means that when that cohort reaches the late teens or early twenties, there are fewer young males in their highest-crime years, and thus less crime.<sup>7</sup> The second mechanism stems from the fact that abortion has a disproportionate effect on the birth of those who are most at risk of engaging in criminal behaviour. Teenagers, unmarried women, and the economically disadvantaged are all substantially more likely to seek abortions (Levine et al., 1999). Recent studies have found children born to these mothers to be at higher risk of committing crime in adolescence (Comanor & Philipps, 2002). Thus, the two mechanisms form a mechanism

<sup>7</sup>It is shown that the legalisation of abortion leads to a drop in birth rates and that the crime rate would be expected to fall accordingly (Levine et al., 1999). For example, consider a town with the population of 10,000. Suppose that the birth rate before the legalisation of abortion is constant at 2%, while that after the legalisation of abortion is 1%. Also assume that the overall crime rate is 5%, and those who are aged 18–24 commit the half of crimes. All other things being equal, one can infer that the overall crime rate will drop by approximately 1% 24 years after the legalisation of abortion.

complex linking the legalisation of abortion in the early 1970s to the drops in crime in the early 1990s.

Donohue and Levitt look for evidence of correlation by focusing on the variations of national time series of crime and abortion, of differential crime patterns between states which legalised abortion early and other states, and of state abortion rates and the state crime rates. They show that the legalisation of abortion was associated with a subsequent drop in crime. All of violent crime, property crime, and murder have fallen steadily since 1991, roughly the time the first cohort born would hit its criminal prime. Additionally, the five states that legalised abortion in 1970 saw drops in crime before the other 45 states and Washington DC, which legalised abortion in 1973. Also, higher rates of abortion in a state in the 1970s and early 1980s are strongly linked to lower crime over the period from 1985 to 1997. Moreover, the observed correlation also holds conditional on various potential confounders, such as such as the level of incarceration, the number of police, and measures of the state's economic well-being (the unemployment rate, income per capita, and poverty rate). It is shown that there is no relationship between abortion rates in the mid-1970s and crime changes between 1972 and 1985, when the cohort directly affected by abortion legislation would have been very young. Almost all of the abortion-related crime decrease can be attributed to reductions in crime among cohorts born after abortion legalisation. In contrast, there is little change in crime among older cohorts, who were not affected by abortion legalisation. The correlation is further supported by the more recent study of Donohue and Levitt (2019).

It is clear that Donohue and Levitt's justification of the causal claim about legalised abortion and crime rates accords well with Evidential Pluralism. Not only do they look for a conditional correlation, but they also seek relevant mechanisms and they show that these mechanisms can account for the extent of the observed correlation. What is more, Donohue and Levitt's criticisms of the alternative causal explanations are also compatible with the epistemological picture provided by Evidential Pluralism. For example, the reason that Donohue and Levitt dismiss factors such as the increasing use of incarceration and the rise in police numbers as the causes of the drop in crime rates is that these trends fail to exhibit an appropriate conditional correlation.

Donohue and Levitt's study has sparked debate and controversy in the literature. Both their evidence of correlation and their evidence of mechanisms have been disputed (Joyce, 2004; Lott & Whitley, 2007; Chamlin et al., 2008; Foote & Goetz, 2008; Dills et al., 2010). Chamlin et al. (2008), for example, argue that there is no evidence that the legalisation of abortion led to a decline in the birth rate for teenage or unmarried women: i.e., they are sceptical of the evidence of mechanisms. Moreover, Lott and Whitley (2007) question whether Donohue and Levitt provide the complete mechanism complex from abortion to crime. As they point out, 'abortion can eliminate unwanted children and can benefit many women, but it can also make other women who are unable to bring themselves to have an abortion worse-off and more likely to have out-of-wedlock births' (Lott & Whitley, 2007, p. 324). Joyce (2004) challenges both correlation and mechanism. In response, Donohue and Levitt (2019) defend their causal claim with updated evidence, which is relevant to both correlation and mechanism. That both sides of this debate focus on evidence of correlation and

evidence of mechanism provides some support for the view that Evidential Pluralism captures the structure of causal inference in the social sciences. Whether or not Donohue and Levitt are correct, the debate indicates that good social science research needs to consider both association studies and mechanistic studies when assessing causal claims.

### 2.3 New foundations for mixed methods research

Mixed methods research is now widespread in the social sciences and such research also considers a variety of evidence when assessing causal claims. It is important not to conflate Evidential Pluralism with mixed methods research: there are substantial differences between the two, as we shall explain. Nevertheless, we shall suggest that Evidential Pluralism can provide new foundations for those variants of mixed methods research that seek to establish causal claims.

Mixed methods research is usually defined as a methodology or a methodological orientation employing both qualitative and quantitative data, methods, or designs:

With mixed methods research, researchers combine elements of qualitative and quantitative research approaches (e.g., use of qualitative and quantitative viewpoints, data collection, analysis, inference techniques) for the purposes of breadth and depth of understanding, and for mutual corroboration (Johnson et al., 2007, p. 123).

[W]e defined mixed-method designs as those that include at least one quantitative method (designed to collect numbers) and one qualitative method (designed to collect words), where neither type of method is inherently linked to any particular inquiry paradigm (Greene et al., 1989, p. 256).

In mixed methods, the researcher (i) collects and analyses both qualitative and quantitative data rigorously in response to research questions and hypotheses; (ii) integrates (or mixes or combines) the two forms of data and their results; (iii) organises these procedures into specific research designs that provide the logic and procedures for conducting the study; and (iv) frames these procedures within theory and philosophy (Creswell & Plano Clark, 2018, p. 5).

Evidential Pluralism, on the other hand, does not appeal to the qualitative / quantitative distinction. Instead it features two other distinctions: (i) the distinction between evidence of the existence of a conditional correlation and evidence of the existence of an appropriate mechanism complex, and (ii) the distinction between association studies and mechanistic studies. It is true that association studies usually use quantitative methods or designs, while mechanistic studies often employ qualitative methods or designs. But this is not always the case. Mechanistic studies can have quantitative elements: e.g., when investigating mechanisms linking legalised abortion and crime rates, Donohue and Levitt (2001) appeal to various quantitative data and methods. Moreover, association studies can include qualitative elements: e.g., when examining the association between economic inequality and democratic transitions, Haggard & Kaufman (2016, pp. 102–141) employ qualitative methods such as process tracing.

It is also important to note that Evidential Pluralism is purely an account of the epistemology of causation, while mixed methods research is a methodology that sometimes invokes metaphysical presuppositions. In the social sciences, a variety of metaphysical stances are invoked to motivate the use of certain types of qualitative or quantitative study: for example, positivism, postpositivism, constructivism, interpretivism, and, as we have seen, critical realism. These metaphysical stances tend to be mutually incompatible, which creates a tension in mixed methods research. Accordingly, the options for the mixed methods researcher are to live with this tension, or to reject any appeal to ‘inquiry paradigms’—see the above quote of Greene et al. (1989, p. 256). Either way, mixed methods research apparently lacks coherent philosophical foundations.

A typical response to this dilemma is to appeal to pragmatism. This response is rooted in American pragmatism, especially the works of John Dewey, Charles Sanders Peirce and Richard Rorty (Cherryholmes, 1992; Johnson & Onwuegbuzie, 2004; Feilzer, 2010; Tebes, 2012), and it highlights the instrumental role of theories in inquiry (Johnson & Onwuegbuzie, 2004; Johnson & Gray, 2010; Morgan, 2014). According to the pragmatist response, both the mind-independent physical world and the constructed social and psychological world exist, and social reality is a product of both; social scientific research is value-oriented; and the aim of social scientific research is to solve problems. The claim is that social scientists do not have to choose between postpositivism and constructivism / interpretivism (Johnson & Onwuegbuzie, 2004; Teddlie & Tashakkori, 2009; Creswell & Plano Clark, 2018). They are free to choose the methods that best meet their needs and purposes and they can employ both quantitative and qualitative methods in their research (Johnson & Onwuegbuzie, 2004; Greene, 2006; Teddlie & Tashakkori, 2009; Creswell & Plano Clark, 2018).

Unfortunately, the pragmatist response provides rather weak philosophical foundations for mixed methods research: it motivates the consideration of mixed methods as one of several possible options, but does not offer any reason to think that mixed methods might be better than a single method on its own. As Greene and Hall (2010, p. 138) observe, ‘Whatever works; whatever can best engage and usefully inform the important practical problem at hand’ is the best methodology, from the pragmatist perspective. And Tashakkori and Teddlie (1998, p. 24) note, ‘Decisions regarding the use of either qualitative or quantitative methods (or both) depend upon the research question.’ In short, the pragmatist position merely justifies the inclusion of mixed methods in a portfolio of possible research designs—it does not provide normative grounds for using mixed methods. Therefore, the pragmatist position provides at best ‘an attractive philosophical partner for mixed methods research,’ as Johnson and Onwuegbuzie (2004, p. 14) put it.

Evidential Pluralism can help by providing normative grounds for using mixed methods. Evidential Pluralism is well motivated as a theory of the epistemology of causality, and it in turn motivates the use of both qualitative and quantitative methods. Quantitative methods employed by association studies can confirm both the existence of an appropriate correlation as well as the existence of an appropriate mechanism (channels  $C_1$  and  $C_2$  in Fig. 1). Quantitative methods can also feature in mechanistic

studies that investigate particular links or features of the mechanism of action (channel  $M_1$ ). Other mechanistic studies can use qualitative methods to test and explore specific mechanism hypotheses (also  $M_1$ ). It is precisely because one should consider both association and mechanistic studies that one should pay attention to both quantitative and qualitative methods.

In addition, Evidential Pluralism can provide guidance on how to integrate quantitative and qualitative methods, which is often a challenge for the mixed methods researcher. Again, Fig. 1 provides the structure of the integration task. One point at which qualitative and quantitative methods need to be integrated is in the assessment of specific mechanism hypotheses: the question is the extent to which key features of relevant mechanisms are confirmed by qualitative and quantitative methods. Another point of integration is in the assessment of the general mechanistic claim that there exists a mechanism complex that explains instances of the putative effect in terms of instances of the putative cause and that can account for the extent of the observed correlation. As we see in Fig. 1, one needs to consider quantitative data from association studies at this stage, as well as the statuses of the specific mechanism hypotheses. Then there is the assessment of the claim that the putative cause and effect are correlated, conditional on any potential confounding variables suggested by background evidence or theory. At this stage, quantitative data will usually be most relevant, although qualitative methods may also have an influence through channel  $M_3$ . Finally, the status of the causal claim depends on the statuses of the correlation claim and the general mechanistic claim, and it is at this point that all the qualitative and quantitative data are integrated. Parkkinen et al. (2018, § 7.1) provide some guidance on all these points of integration.

We see then that Evidential Pluralism can help to justify the use of mixed methods in causal analysis and to structure the data integration task. It is in this sense that Evidential Pluralism can help to provide coherent philosophical foundations for mixed methods research.<sup>8</sup> These foundations can be viewed as complementary to those provided by other approaches, such as pragmatism, or can act in a standalone capacity.

Weinstein's study of insurgent violence is a good example of the use of mixed methods research to establish correlation and mechanism. Weinstein (2007) proposes a theory to explain differences between the ways in which rebel groups employ violence. He argues that resources and financing are one key causal factor of the strategies of violence: 'rebel groups that emerge in environments rich in natural resources or with the external support of an outside patron tend to commit high levels of indiscriminate violence; movements that arise in resource-poor contexts perpetrate far fewer abuses and employ violence selectively and strategically' (Weinstein, 2007, p. 7). Weinstein's theory not only predicts the correlation between the initial endowment to which rebel leaders have access and their use of violence, but also identifies some mechanisms linking them. For instance, Weinstein

---

<sup>8</sup>This justification applies where mixed methods research is used to establish or assess causal claims. Qualitative and quantitative approaches can also be combined for other purposes: e.g., for hypothesis generation, to gain a general understanding of the social context, or to ensure that certain stakeholders have a voice that is represented in the evidence base. Evidential Pluralism does not speak to these other tasks.

argues that resources shape the membership profile of a rebel group, which in turn affects its internal organisation and strategies it uses in war. In order to support his theory, Weinstein integrates qualitative interview-based studies of the rebel groups and community-level social histories with statistical analysis of original newspaper datasets on patterns of violence in four case studies of rebel groups in Mozambique, Peru, and Uganda. The quantitative data provides solid evidence of correlation that supports his theory at the national level, while the qualitative data offers more nuanced mechanistic evidence at the subnational level. Note that Weinstein employs mixed methods both to test the general causal claim and also to assess specific mechanistic hypotheses.<sup>9</sup>

Another example is Ivankova and Stick's study of PhD students' persistence. Ivankova and Stick (2007) investigate factors that contribute to students' persistence in a remote-learning doctoral programme. In order to identify factors, they use a two-phase study, starting with a quantitative approach and followed by a qualitative approach. The roles of the two approaches are clearly stated as follows:

In this study, the quantitative data helped identify a potential predictive power of selected external and internal factors on the distributed doctoral students' persistence and purposefully select the informants for the second phase. Then, a qualitative multiple case study approach was used to explain why certain external and internal factors, tested in the first phase, were significant predictors of students' persistence in the program. Thus, the quantitative data and results provided a general picture of the research problem, while the qualitative data and its analysis refined and explained those statistical results by exploring the participants' views regarding their persistence in more depth. (Ivankova & Stick, 2007, p. 97.)

Interestingly Ivankova and Stick largely avoid causal terminology—perhaps influenced by the legacy of Yule and other early social scientists. They explicitly classify their study design as a 'sequential explanatory mixed methods design', i.e., as first identifying factors associated with persistence and then finding explanations of these associations. But it is apparent that their task is really causal: they use their results to make a series of recommendations for how to improve persistence in such programmes, and this move would only be warranted if the factors they have identified are causes—rather than merely correlates—of persistence.

From the point of view of Evidential Pluralism, there is no need for any reticence here with regard to causal claims. This is because Ivankova and Stick have done what they need to do to confirm causality. They used a quantitative association study to identify factors correlated with persistence in the programme and then used a qualitative mechanistic study to provide evidence that there are mechanisms that invoke these factors to explain persistence (or drop-out). Taken together, these

---

<sup>9</sup>For example, a combination of interview-based studies and statistical analysis is used to confirm the causal connection between Renamo's indiscriminate abuse on non-combatant populations and the full financial and military backing of the Rhodesian government. For more detail, see Weinstein (2007, pp. 229–234).

studies provide some good evidence for causality.<sup>10</sup> In situations such as this, an appeal to Evidential Pluralism might give researchers the confidence to draw causal conclusions.

In sum, Evidential Pluralism can provide coherent philosophical foundations for mixed methods research as applied to causal inquiry. It can also provide guidance on how to integrate these quantitative and qualitative methods.

### 3 Objection: are correlation and mechanism insufficient?

Having presented the case for applying Evidential Pluralism to the social sciences, we now address two objections: an objection to the claim that establishing correlation and mechanism is sufficient for establishing causation, which we tackle in this section, and an objection to the claim that establishing both correlation and mechanism is necessary for establishing causation, which we address in the next section. Understanding how it is immune to these objections will help to highlight some important features of Evidential Pluralism.

One way to object to Evidential Pluralism is to criticise the top part of Fig. 1, namely the thesis that establishing correlation and mechanism is sufficient to establish causality. Reiss (2009, § 4) adopts this strategy, for instance.<sup>11</sup> He employs the following hypothetical example. Suppose association studies establish a correlation between watching violent TV ( $W$ ) and violent behaviour ( $V$ ). Given that a correlation on its own fails to establish the causal claim that  $W$  causes  $V$ , Evidential Pluralism would suggest that one should also consider evidence of mechanisms. Now suppose that some such evidence does establish that there is a mechanism by which watching violent TV leads to increased aggression, which in turn leads to violent behaviour. (For example, there may be a psychological mechanism according to which viewers identify with the aggressive characters on TV and think of the depicted scenarios as realistic, which then results in more violent behaviour in real life.) One might then conclude that  $W$  is a cause of  $V$ , since both a correlation and a mechanism are established. But this conclusion may well be erroneous. There may be a masking mechanism by which  $W$  acts as a deterrent in some individuals, reducing  $V$ . Suppose these two mechanisms exactly cancel out and that the correlation between  $W$  and  $V$  is in fact attributable to an unmeasured confounder—socio-economic status, say. Then  $W$  is not a cause of  $V$  after all, and Evidential Pluralism seems to have led us astray.

---

<sup>10</sup>Whether they have done enough to *establish* causality is not a question that we shall attempt to settle here.

<sup>11</sup>Note that Reiss is a conceptual pluralist about causation: he thinks there are multiple meanings of the word ‘cause’, and that evidence of mechanisms and evidence of correlation may latch on to different concepts of cause. This is a view that we do not subscribe to, for reasons espoused by Williamson (2006, 2013), and it is worth emphasising that Evidential Pluralism need not be accompanied by either conceptual pluralism (the view that we employ different concepts of cause on different occasions) or metaphysical pluralism (the view that there are multiple causal relations in reality). Although Reiss presents the following objection with an appeal to conceptual pluralism, his objection also applies to the more standard view that evidence of mechanisms and evidence of correlation are evidence for the same concept of cause.

There are three points to make in response to this objection. First, as we elaborate in Section 1, Evidential Pluralism requires establishing a correlation conditional on potential confounders, not an unconditional correlation. Thus in Reiss' example, correlation in the appropriate sense is not established, after all. Hence, nor is the causal claim. Socio-economic status is an obvious potential confounder, so one would need to test for its influence.

Second, as we emphasise in Section 1, one needs to establish that the *complex of mechanisms* linking the putative cause to the putative effect can account for the observed correlation. It is not enough to consider a single pathway of action: one needs to rule out masking by counteracting mechanisms. *If* it is established that there is a mechanism of action and *if* potential counteracting mechanisms can be shown to have negligible influence and *if* the observed correlation is large enough to infer that the putative cause and effect are genuinely correlated conditional on potential confounders, *then* one is in a position to establish the causal claim. The larger the conditional correlation, the less likely it is that there remain unmeasured confounding variables and unconsidered counteracting mechanisms: if they were responsible for a large correlation, it would be likely that we would know about them and have marked them down as potential confounders / counteractors. A psychological mechanism by which watching violence discourages violence is an obvious potential counteracting mechanism, so one would need to examine its influence.

Thus Reiss' objection is really an objection to a caricature of Evidential Pluralism, rather than to Evidential Pluralism itself. But there is a third point to make about Reiss' objection, which highlights an important aspect of Evidential Pluralism: Evidential Pluralism is an account of the epistemology of causality, not an analysis of causality. If it were an analysis of causality of the form *A causes B if and only if A and B are appropriately correlated and linked by an appropriate mechanism complex* then a single hypothetical counterexample would refute it. But Evidential Pluralism is an account of the assessment of causal claims, rather than an analysis of causality. Establishing a causal claim is a fallible activity; that one might sometimes be mistaken about whether *A causes B* does not refute the thesis that the best method for establishing that *A causes B* requires establishing that *A* and *B* are appropriately correlated and linked by an appropriate mechanism complex. As long as errors are sufficiently rare, this method may yet lead to the optimal progression of science. Indeed, requiring immunity from all possible doubt when establishing a proposition would prevent that proposition from ever being established. We must balance the need to avoid falsity with the need to establish truths.

#### 4 Objection: is correlation unnecessary?

We now turn to a second objection to the application of Evidential Pluralism to the social sciences: an objection to the claim that establishing both correlation and mechanism is necessary for establishing causation.

In particular, the necessity of establishing correlation has been challenged. Claveau (2012) considers a counterexample drawn from economics and argues that the following causal claim was established without establishing correlation:

- (C) The strictness of unemployment benefit eligibility ( $S$ ) reduces the unemployment rate ( $U$ ).

Claveau suggests that there is a consensus amongst economists that a monitoring and sanctions system for unemployment benefits has a powerful effect on the unemployment rate. Various studies have sought to show that benefit sanctions induce a sharp increase in the exit rate from unemployment to employment across different countries. In other words, the strictness of unemployment benefit eligibility is a negative cause of the unemployment rate across countries. Claveau (2012) suggests that this causal claim was established by mechanistic evidence alone:

The reason why difference-making evidence was not relied on is rather trivial: there was no measure of [ $S$ ] comparable across countries (and there is still none as far as I am aware). Economists drew on the available evidence, i.e. mechanistic evidence. And it seems that this evidence—including clear model predictions and micro-data evidence in line with them . . .—was sufficient to gather general support to the claim. (Claveau, 2012, p. 812.)

Claveau's argument rests on two claims: (i) there is no universal quantitative measure of the strictness of the monitoring and sanctions; (ii) there is only mechanistic evidence for  $C$ . Because of (i), Claveau argues that a correlation between  $S$  and  $U$  across countries is difficult to establish. And thus, given (ii) and given that  $C$  is established, it can be concluded that establishing the existence of a correlation is not required to establish a causal claim.

We would resist Claveau's conclusions as follows.

Firstly, there was good evidence of correlation in this case. That there was no good universal measure of  $S$  readily available does not imply that  $S$  is unmeasurable and that there is no way of establishing correlation. Claveau (2012) challenges the use of a coefficient associated with the overall cost of 'active labour market policies' (ALMPs) as a good universal measure across the countries. He argues that ALMPs include many factors which are unrelated to  $S$ , such as placement services, subsidised training, and subsidised employment. However, this does not eliminate the possibility of a good measure of  $S$  in each particular country, even if a good universal measure is not to hand. Moreover, local measures of  $S$  can be used to establish a correlation with  $U$  that holds across countries. For example, van den Berg et al. (2004) show that the imposition of a particular sanction, namely additional job search assistance, is significantly positively associated with the transition rate from welfare to work in the Netherlands, while Lalive et al. (2005) show that warning about benefits reduction is positively correlated with the transition rate in Switzerland. These studies provide useful evidence of a correlation between  $U$  and (some realisation of)  $S$  across the Netherlands and Switzerland. Thus, a correlation can be established across countries even if (i) holds.

Second, in this case the mechanistic evidence was arguably not enough on its own to establish causation generally. At best, the existence of a mechanism from benefit sanctions to employment was established in the USA, UK, the Netherlands, and Switzerland (Johnson & Klepinger, 1994; Dolton & O'Neill, 1996; van den Berg

et al., 2004; Lalive et al., 2005). More would need to be done to establish a general mechanistic claim that holds more widely across countries: it would need to be shown that the mechanisms are extrapolable to other countries. Thus it appears that the causal claim has not been established generally, but only in specific countries. It seems to be too hasty to infer that  $S$  reduces  $U$  generally.

In sum, it appears, contra Claveau, that claim  $C$  may not have been generally established, and that correlation was established in those countries in which  $C$  was established (with the help of mechanistic studies, as per channel  $M_3$  in Fig. 1). Thus Evidential Pluralism fits this case after all.<sup>12</sup>

One might think that more convincing counterexamples to the necessity of establishing correlation emerge in the context of political science. Indeed, political scientists often make within-case (i.e., single-case) causal inferences by employing process-tracing methods, and these methods apparently only seek to identify mechanisms (Bennett, 2010, p. 214). If so, this would cast doubt on the need to establish evidence of correlation when establishing a causal claim.

However, such cases are not genuine counterexamples: in those process-tracing studies used for causal inference, evidence of correlation is present, albeit often implicit, as we shall now see.

According to Beach and Pedersen (2013), process-tracing methods are mainly used in two situations:

- (I) when we know that a correlation exists between  $X$  and  $Y$  but we are in the dark regarding potential mechanisms linking the two ( $X$ - $Y$ -centric theory-building) as we have no theory to guide us; or (II) when we know an outcome ( $Y$ ) but are unsure about the causes ( $Y$ -centric theory building). (Beach & Pedersen, 2013, p. 16.)

Neither situation provides a counterexample to Evidential Pluralism. In  $X$ - $Y$ -centric situations, correlation has already been established: ‘we know that a correlation exists between  $X$  and  $Y$ .’ In  $Y$ -centric situations, evidence of correlation is also required. As Beach and Pedersen (2013, p. 16) note, ‘in the [ $Y$ -centric] instance, the analysis first traces backward from  $Y$  to uncover a plausible  $X$ , turning the study into an  $X$ - $Y$ -Centric analysis.’ Thus, in order to utilise process-tracing method in the  $Y$ -centric situation, one needs to turn a  $Y$ -centric analysis to an  $X$ - $Y$ -centric analysis, which assumes a correlation between  $X$  and  $Y$ . In other words, in uncovering a plausible  $X$ , one searches for a factor for which correlation with  $Y$  has already been established. So, as Bennett (2010, pp. 208–209) also observes, process tracing is mainly used to provide further evidence for causation, having already obtained evidence of correlation.

---

<sup>12</sup>It is worth noting that Claveau’s ‘mechanistic evidence’ does not coincide with what we call ‘mechanistic studies.’ For Claveau (2012, p. 810), ‘[t]he first step in getting to mechanistic evidence is to redescribe the two relata [i.e., cause and effect] at a lower level’. In the case of unemployment benefits and unemployment rates, Claveau takes mechanistic evidence for  $C$  to be evidence that  $S$  reduces  $U$  in the particular countries, say, the USA. However, such evidence would not on its own constitute a mechanistic study, as it does not shed light on features of a mechanism linking  $S$  to  $U$ .

Beach and Pedersen (2013) also indicate a third situation in which process tracing is used: this is where it is used for the purpose of explaining a particular outcome or event. One might think that these explaining-outcome process-tracing studies pose a problem for Evidential Pluralism, as correlation is rarely explicitly discussed in such studies. In response, we would make the following two points.

On the one hand, many explaining-outcome process tracing studies do not seek to establish causal claims. Rather, they aim to ‘explain particular historical outcomes’ (Beach & Pedersen, 2013, p. 63) and historical explanations are not always causal. For example, Wood (2003) is cited by Beach and Pedersen (2013, p. 63) as an example of explaining-outcome process-tracing. Wood provides an explanation of insurgent collective action in El Salvador by identifying ‘three reasons that participants supported the mobilisation and insurgency’ (Wood, 2003, p. 231). This study stops short of establishing causal claims. For this reason, many explaining-outcome process-tracing studies fall outside the remit of Evidential Pluralism.

On the other hand, Evidential Pluralism does indeed apply to situations in which explaining-outcome studies seek specifically *causal* explanations. Moreover we contend that such studies require evidence of correlation. As Beach and Pedersen (2013, p. 18) observe, where explaining-outcome process-tracing studies seek causes, they aim ‘to craft a minimally sufficient explanation of a particular outcome, with sufficiency defined as an explanation that accounts for all of the important aspects of an outcome with no redundant parts being present.’ In order to establish  $X$  as a minimally sufficient explanation of  $Y$  by process tracing, there needs to be a correlation between  $X$  and  $Y$ , for otherwise  $X$  would give no reason to expect  $Y$ . Here the causal claim is single-case (i.e.,  $X$  and  $Y$  are single-case outcomes), so the correlation required is a single-case correlation: the chance of  $Y$  given  $X$  differs from the chance of  $Y$  in the absence of  $X$ , conditional on potential confounders. Evidence for this single-case correlation may take the form of association studies that support a generic correlation in a population to which this particular case belongs. But the evidence might also include mechanistic studies that elucidate the mechanisms in operation in the presence of  $X$  and those in operation in the absence of  $X$ . A comparison of these two cases can sometimes confirm correlation via routes  $M_1$  and  $M_3$  of Fig. 1.<sup>13</sup> Note that Evidential Pluralism does not require that any of these studies be new studies—the correlation may have been established previously.

Skocpol’s study on social revolutions includes examples of this kind of process tracing (Goldstone, 1997; George & Bennett, 2005; Mahoney, 2012). Skocpol (1979)

---

<sup>13</sup>For example, a comparison between the mechanisms operating when a parachute is engaged and those operating when no parachute is engaged provide some evidence of a correlation between a particular parachute use and serious injury (see Section 1).

Thus, there is a distinction between establishing a case-specific mechanism from the presence of  $X$  to  $Y$  and establishing that  $X$  is a cause of  $Y$ , because the latter requires establishing correlation, which depends on a comparison with the case in which  $X$  is absent. For example, having established a case-specific mechanism between the assassination of Archduke Ferdinand and the outbreak of the First World War, further evidence of correlation is needed in order to establish the claim that the assassination of Archduke Ferdinand is a cause of the outbreak of the First World War. This further evidence would need to establish that the chance of the outbreak of World War I would have differed in the absence of the assassination.

argues that international pressure and peasant rebellion are two causes of social revolutions in France, Russia, and China. She uses process tracing to show how each of the two independent variables (i.e., international pressure and peasant rebellion) set into motion a complex sequence of events that culminate in revolutionary social transformation in each country. Overall, as George and Bennett (2005, p. 227) argue, Skocpol's use of process tracing shows 'how these two variables were causally related to the revolutionary social transformation in each of these countries.' Skocpol does this by providing evidence of correlation in addition to evidence of mechanism: by unpacking the detail of the mechanisms and carrying out a contrastive analysis, Skocpol argues that international pressure and peasant rebellion are associated with outbreaks of social revolutions.<sup>14</sup>

Skocpol can be thought of as providing examples of explaining-outcome process tracing for the following reason. Beach and Pedersen (2013, p. 19) construe explaining-outcome process tracing as an iterative research strategy that aims to trace 'case-specific mechanisms' that 'cannot be detached from the particular case'. However, 'explaining-outcome studies often have theoretical ambitions that reach beyond the single case' (Beach & Pedersen, 2013, p. 19). These descriptions seem to fit Skocpol's work well, as she identifies three case-specific mechanisms and connects them to two causal variables. Thus her study has ambitions to connect the three cases, but she stops short of making any generic causal claim.

We see, then, that evidence of correlation is required even where political scientists make within-case causal inferences by employing process-tracing methods. A consideration of these within-case inferences shows that Evidential Pluralism offers more than an account of the epistemology of generic (i.e., repeatedly instantiable) causal claims in the social sciences—it also provides an account of the epistemology of single-case causal claims.

## 5 Conclusions

In this paper we set out a specific approach to the assessment of causal claims, Evidential Pluralism, and argued that this approach can be fruitfully applied to the social sciences. In particular, we argued that applying Evidential Pluralism to the social sciences yields the following benefits. Firstly, it motivates a new methodology for evidence-based policy. This methodology extends the evidence base to include mechanistic studies, which are undervalued by present-day EBP. Second, it provides an account of the confirmatory relationships in basic social science research. This can help researchers to structure their own efforts to establish causal claims, as well as

---

<sup>14</sup>Skocpol's contrastive analysis involves comparing the abortive Russian Revolution of 1905 with the successful Russian Revolution of 1917 and using some aspects of English, Japanese, and German history as contrasts to those of French, Russian, and Chinese history.

A question arises as to whether this contrastive analysis, which can be thought of as an instance of Mill's method of difference, is a process-tracing technique or something that goes beyond process tracing (Goldstone, 1997; George & Bennett, 2005). Either way, Skocpol's study clearly provides both evidence of correlation and evidence of mechanism for her within-case causal inferences.

to understand and evaluate the attempts of others. Third, Evidential Pluralism can help to provide coherent philosophical foundations for mixed methods research: it motivates the need to integrate quantitative and qualitative research in the context of causal analysis. We also defended Evidential Pluralism against two objections that stem from social science examples—objections to the necessity and sufficiency of evidence of correlation and evidence of mechanisms—and saw that Evidential Pluralism offers an account of the epistemology of both generic and single-case causal claims.

Two key points are worth reiterating.

Firstly, Evidential Pluralism holds that in order to establish a causal claim, one normally needs to establish both the existence of an appropriate conditional correlation and the existence of an appropriate mechanism complex. Association studies and mechanistic studies can provide evidence that is relevant to both these latter claims: it should not be assumed that association studies only provide evidence of correlation, nor that mechanistic studies only provide evidence of mechanisms.

Second, Evidential Pluralism is a normative epistemological thesis, concerning how one ought to assess causal claims. We do not maintain that every actual assessment of a causal claim in the social sciences can be rationally reconstructed as conforming to Evidential Pluralism. In the social sciences, not all science is good science. Nevertheless, we do hold that episodes of good science do conform to Evidential Pluralism, and that Evidential Pluralism can provide a useful tool for research scientists and policy makers to help structure research and evaluation. This is not to say that assessing a causal claim in accordance with Evidential Pluralism will guarantee success. The act of establishing a causal claim is fallible, and even good science can be misled, where evidence is misleading.

Finally, we acknowledge that a programmatic paper like this is liable to raise more questions than it has space to answer. For example, the precise relationship between Evidential Pluralism and related methods such as process tracing and contribution analysis merits a more detailed exploration than is possible here. Another important task for further research would be to situate Evidential Pluralism more precisely within the rich plethora of philosophical traditions in the social sciences. It would also be interesting to investigate whether Evidential Pluralism can be extrapolated beyond the biomedical and social sciences—e.g., to the physical sciences.

**Acknowledgements** We are grateful to the Leverhulme Trust for funding this research (grant RPG2019-059), and to Ben Baumberg Geiger, Virginia Ghiara, Samuel D. Taylor, Erik Weber and the anonymous referees for many helpful suggestions.

**Funding** The Leverhulme Trust Research Project Grant (grant RPG-2019-059).

## Declarations

**Conflict of Interests** The authors declare no conflict of interest.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons

licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Baron, J. (2018). A brief history of evidence-based policy. *The ANNALS of the American Academy of Political and Social Science*, 678(1), 40–50.
- Beach, D., & Pedersen, R. B. (2013). *Process-tracing methods*. The University of Michigan Press.
- Bennett, A. (2010). Process tracing and causal inference. In H. E. Brady, D. Collier, & A. Bennett (Eds.) *Rethinking social inquiry*. 2nd edn. (pp. 207–220). Lanham Rowman and Littlefield.
- Bonell, C., Fletcher, A., Morton, M., Lorenc, T., & Moore, L. (2012). Realist randomised controlled trials: A new approach to evaluating complex public health interventions. *Social Science & Medicine*, 75(12), 2299–2306. Special Issue: Place, migration & health.
- Cabinet Office (2018). What Works Network: Membership criteria. [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/747965/WW-membership-paper-Oct2018.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/747965/WW-membership-paper-Oct2018.pdf)
- Cartwright, N., & Hardie, J. (2012). *Evidence-based policy: A practical guide to doing it better*. Oxford University Press.
- Chamlin, M. B., Myer, A. J., Sanders, B. A., & Cochran, J.K. (2008). Abortion as crime control: A cautionary tale. *Criminal Justice Policy Review*, 19, 135–152.
- Cherryholmes, C. H. (1992). Notes on pragmatism and scientific realism. *Educational Researcher*, 21(6), 13–17.
- Claveau, F. (2012). The Russo-Williamson theses in the social sciences: Causal inference drawing on two types of evidence. *Studies in History and Philosophy of Biological and Biomedical Sciences*, 43(4), 806–813.
- Comanor, W. S., & Philipps, L. (2002). The impact of income and family structure on delinquency. *Journal of Applied Economics*, 5(2), 209–232.
- Copeland, D. C. (2015). *Economic interdependence and war*. Princeton University Press.
- Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research*. SAGE.
- Demeulenaere, P. (ed.) (2011). *Analytical sociology and social mechanisms*. Cambridge University Press.
- Dills, A. K., Miron, J. A., & Summers, G. (2010). What do economists know about crime? In R.D. Tella, S. Edwards, & E. Schargrodsy (Eds.), *The economics of crime: Lessons for and from Latin American* (pp. 269–302). University of Chicago Press.
- Dolton, P., & O'Neill, D. (1996). Unemployment and the restart effect: Some experimental evidence. *Economic Journal*, 106(435), 387–400.
- Donohue, J. J., & Levitt, S. D. (2019). The impact of legalized abortion on crime over the last two decades. Working Paper 25863, National Bureau of Economic Research.
- Donohue, J. J., & Levitt, S. D. (2001). The impact of legalized abortion on crime. *The Quarterly Journal of Economics*, 116(2), 379–420.
- Feilzer, M. Y. (2010). Doing mixed methods research pragmatically: Implications for the rediscovery of pragmatism as a research paradigm. *Journal of Mixed Methods Research*, 4(1), 6–16.
- Foote, C. L., & Goetz, C. F. (2008). The impact of legalized abortion on crime: Comment. *The Quarterly Journal of Economics*, 123, 407–423.
- George, A. L., & Bennett, A. (2005). *Case studies and theory development in the social sciences*. The MIT Press.
- Ghiara, V., & Russo, F. (2019). Reconstructing the mixed mechanisms of health: The role of bio- and sociomarkers. *Longitudinal and Life Course Studies*, 10(1), 7–25.
- Goldstone, J. A. (1997). Methodological issues in comparative macrosociology. *Comparative Social Research*, 16, 107–120.
- Greene, J. C. (2006). Towards a methodology of mixed methods social inquiry. *Research in the Schools*, 13(1), 93–98.
- Greene, J. C., Caracelli, V. J., & Graham, W.F. (1989). Toward a conceptual framework for mixed-method evaluation designs. *Educational Evaluation and Policy Analysis*, 11(3), 255–274.

- Greene, J. C., & Hall, J. N. (2010). Dialectics and pragmatism: Being of consequence. In A. Tashakkori, & C. Teddlie (Eds.), *Sage handbook of mixed methods in social and behavioral research*. 2nd edn. (pp. 119–144). SAGE Publications, Inc.
- Haggard, S., & Kaufman, R. R. (2016). *Dictators and democrats*. Princeton University Press.
- Hedström, P., & Swedberg, R. (1998). Social mechanisms: An introductory essay. *Social mechanisms: An analytical approach to social theory* (pp. 1–31). Cambridge University Press.
- Hedström, P., & Ylikoski, P. (2010). Causal mechanisms in the social sciences. *The Annual Review of Sociology*, 36, 49–67.
- Illari, P. M. (2011). Disambiguating the Russo-Williamson thesis. *International Studies in the Philosophy of Science*, 25(2), 139–157.
- Illari, P. M., & Williamson, J. (2012). What is a mechanism? Thinking about mechanisms across the sciences. *European Journal for Philosophy of Science*, 2, 119–135.
- Ioannidis, S., & Psillos, S. (2017). In defense of methodological mechanism: The case of apoptosis. *Axiomathes*, 27(6), 601–619.
- Ivankova, N. V., & Stick, S. L. (2007). Students' persistence in a distributed doctoral program in educational leadership in higher education: A mixed methods study. *Research in Higher Education*, 48(1), 93–135.
- Jagosh, J., Tilley, N., & Stern, E. (2016). Realist evaluation at 25: Cumulating knowledge, advancing debates and innovating methods. *Evaluation*, 22(3), 267–269.
- Johnson, R. B., & Gray, R. (2010). A history of philosophical and theoretical issues for mixed methods research. In A. Tashakkori, & C. Teddlie (Eds.), *Sage Handbook of Mixed Methods in Social and Behavioral Research* (2nd edn, pp. 6994). SAGE Publications, Inc.
- Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Research*, 33(7), 14–26.
- Johnson, R. B., Onwuegbuzie, A. J., & Turner, L.A. (2007). Toward a definition of mixed methods research. *Journal of Mixed Methods Research*, 1(2), 112–133.
- Johnson, S. D., Tilley, N., & Bowers, K.J. (2015). Introducing EMMIE: An evidence rating scale to encourage mixed-method crime prevention synthesis reviews. *Journal of Experimental Criminology*, 11(3), 459–473.
- Johnson, T., & Klepinger, D. (1994). Experimental evidence on unemployment insurance work-search policies. *Journal of Human Resources*, 29(3), 695–717.
- Joyce, T. (2004). Further tests of abortion and crime. Working Paper 10564, National Bureau of Economic Research.
- Lalive, R., Zweimüller, J., & van Ours, J.C. (2005). The effect of benefit sanctions on the duration of unemployment. *Journal of the European Economic Association*, 3(6), 1386–1417.
- Levine, P., D. S., Kane, T., & Zimmerman, D. (1999). Roe v Wade and American fertility. *American Journal of Public Health*, 89(2), 199–203.
- Lott, J. R., & Whitley, J. (2007). Abortion and crime: Unwanted children and out-of-wedlock births. *Economic Inquiry*, 45, 304–324.
- Machamer, P., Darden, L., & Craver, C. (2000). Thinking about mechanisms. *Philosophy of Science*, 67, 1–25.
- Mahoney, J. (2012). The logic of process tracing tests in the social sciences. *Sociological Methods & Research*, 41(4), 570–597.
- Marchionni, C., & Reijula, S. (2019). What is mechanistic evidence, and why do we need it for evidence-based policy? *Studies in History and Philosophy of Science, Part A*, 73, 54–63.
- Mayne, J. (2001). Addressing attribution through contribution analysis: Using performance measures sensibly. *The Canadian Journal of Program Evaluation*, 16(1), 1–24.
- Moneta, A., & Russo, F. (2014). Causal models and evidential pluralism in econometrics. *Journal of Economic Methodology*, 21(1), 54–76.
- Morgan, D. L. (2014). Pragmatism as a paradigm for social research. *Qualitative Inquiry*, 20(8), 1045–1053.
- Parkkinen, V.-P., Wallmann, C., Wilde, M., Clarke, B., Illari, P., Kelly, M. P., Norell, C., Russo, F., Shaw, B., & Williamson, J. (2018). *Evaluating evidence of mechanisms in medicine: Principles and procedures*. Springer.
- Pawson, R., & Tilley, N. (1997). *Realistic evaluation*. SAGE.
- Reichenbach, H. (1956). *The direction of time* (1971 edn.). The University of California Press.
- Reiss, J. (2009). Causation in the social sciences: Evidence, inference, and purpose. *Philosophy of the Social Sciences*, 39, 20–40.
- Russo, F., & Williamson, J. (2007). Interpreting causality in the health sciences. *International Studies in the Philosophy of Science*, 21(2), 157–170.
- Salmon, W. C. (1998). *Causality and explanation*. Oxford University Press.

- Sewell, W. H., Haller, A. O., & Portes, A. (1969). The educational and early occupational attainment process. *American Sociological Review*, 34(1), 82–92.
- Skocpol, T. (1979). *States and social revolutions: A comparative analysis of France, Russia and China*. Cambridge University Press.
- Sober, E. (2001). Venetian sea levels, British bread prices, and the principle of the common cause. *British Journal for the Philosophy of Science*, 52, 331–346.
- Steel, D. (2008). *Across the boundaries extrapolation in biology and social science*. Oxford University Press.
- Tashakkori, A., & Teddlie, C. (1998). *Mixed methodology: Combining qualitative and quantitative approaches*. Sage.
- Tebes, J. K. (2012). Philosophical foundations of mixed methods research: Implications for research practice. In L. A. Jason, & D. S. Glenwick (Eds.) *Methodological approaches to community-based research* (pp. 13–31). American Psychological Association.
- Teddlie, C., & Tashakkori, A. (2009). *Foundations of mixed methods research*. Sage.
- Thornton, A., Sidebottom, A., Belur, J., Tompson, L., & Bowers, K. (2019). On the development and application of EMMIE: Insights from the what works centre for crime reduction. *Policing and Society*, 29(3), 266–282.
- Tilley, N. (2016). EMMIE and engineering: What works as evidence to improve decisions? *Evaluation*, 22(3), 304–322.
- United Nations (2013). Effective use of statistics in evidence-based policymaking: Conceptual framework. Technical Report E/ESCWA/SD/2013/Technical Paper.1, United Nations Economic and Social Commission for Western Asia, New York. [https://www.unescwa.org/sites/www.unescwa.org/files/publications/files/e\\_escwa\\_sd\\_13\\_tp-1\\_e.pdf](https://www.unescwa.org/sites/www.unescwa.org/files/publications/files/e_escwa_sd_13_tp-1_e.pdf)
- van den Berg, G. J., van der Klaauw, B., & van Ours, J.C. (2004). Punitive sanctions and the transition rate from welfare to work. *Journal of Labor Economics*, 22(1), 211–241.
- Vigen, T. (2015). *Spurious Correlations*. Hachette Books.
- Weber, E. (2009). How probabilistic causation can account for the use of mechanistic evidence. *International Studies in the Philosophy of Science*, 23(3), 277–295.
- Weinstein, J. M. (2007). *Inside rebellion: The politics of insurgent violence*. Cambridge University Press.
- What Works (2018a). The rise of experimental government: Cross-government trial advice panel update report. [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/753468/RiseExperimentalGovernment\\_Cross-GovTrialAdvicePanelUpdateReport.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/753468/RiseExperimentalGovernment_Cross-GovTrialAdvicePanelUpdateReport.pdf). Produced by the What Works Network and the UK Economic and Social Research Council.
- What Works (2018b). The What Works Network: Five years on. [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/677478/6.4154\\_What\\_works\\_report\\_Final.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/677478/6.4154_What_works_report_Final.pdf). Produced 1030 by the What Works Network and the UK Economic and Social Research Council.
- Wilde, M., & Parkkinen, V.-P. (2019). Extrapolation and the Russo-Williamson thesis. *Synthese*, 196, 3251–3262.
- Williamson, J. (2006). Causal pluralism versus epistemic causality. *Philosophica*, 77, 69–96.
- Williamson, J. (2013). How can causal explanations explain? *Erkenntnis*, 78, 257–275.
- Williamson, J. (2019a). Establishing causal claims in medicine. *International Studies in the Philosophy of Science*, 32(2), 33–61.
- Williamson, J. (2019b). Evidential Proximity, Independence, and the evaluation of carcinogenicity. *Journal of Evaluation in Clinical Practice*, 25(6), 955–961.
- Williamson, J. (2021a). A Bayesian account of establishing. *British Journal for the Philosophy of Science*, in press. <https://doi.org/10.1086/714798>
- Williamson, J. (2021b). Calibration for epistemic causality. *Erkenntnis*, 86(4), 941–960.
- Williamson, J. (2021c). Establishing the teratogenicity of Zika and evaluating causal criteria. *Synthese*, 198(10), 2505–2518.
- Williamson, J. (2021d). The feasibility and malleability of EBM+. *THEORIA. An International Journal for Theory History and Foundations of Science*, 36(2), 191–209.
- Wood, E. J. (2003). *Insurgent collective action and civil war in El Salvador*. Cambridge University Press.
- Yamey, G., & Feachem, R. (2011). Evidence-based policymaking in global health—the payoffs and pitfalls. *Evidence Based Medicine*, 16, 97–99.
- Yule, G. U. (1895). On the correlation of total pauperism with proportion of out-relief. *The Economic Journal*, 5(20), 603–611.