Epistemic causality and its application to the social and cognitive sciences

Yafeng Shan, Samuel Taylor and Jon Williamson
Department of Philosophy and Centre for Reasoning
University of Kent, Canterbury, UK

Draft of July 7, 2023

Abstract

The epistemic theory of causality views causality as a tool that helps us to predict, explain and control our world, rather than as a relation that exists independently of our epistemic practices. In this chapter, we first provide an introduction to the epistemic theory of causality. We then outline four considerations that motivate the epistemic theory: the failure of standard theories of causality; parsimony; the epistemology of causality; and neutrality. We illustrate these four considerations in the contexts of the social sciences and the cognitive sciences. We argue that the epistemic theory provides a very natural account of causality across these contexts.

1 Introduction

Just as the epistemic theory of probability (i.e., Bayesianism) interprets probability as a kind of belief, namely rational probabilistic belief, the epistemic theory of causality interprets causality as a kind of belief—rational causal belief, in this case. The epistemic theory was put forward by Williamson (2005, Chapter 9) and further discussed and developed by Williamson (2006a,b, 2007, 2009, 2011, 2013, 2021), Russo (2009, §7.5), Russo and Williamson (2007, 2011a,b), Wilde and Williamson (2016) and Taylor (2021). This chapter provides a brief introduction to epistemic causality in §2, and then motivates the theory from a general point of view and by examining scientific practice.

In §3 we sketch four arguments for epistemic causality: an argument from the failure of more standard approaches (the argument from *failure*, §3.1), an argument that stems from the desideratum of parsimony (the argument from *parsimony* §3.2), an argument that stems from a particular view of the epistemology of causality (the argument from *Evidential Pluralism* §3.3), and an argument from the need to remain neutral about certain questions (the argument from *neutrality*, §3.4).

In §4 we argue that the epistemic theory provides a natural interpretation of causality in the social sciences, illustrating the arguments from failure, parsimony and Evidential Pluralism. In §5 we show that epistemic causality is also well suited to cognitive science, by appeal the arguments from Evidential

Pluralism and neutrality. We conclude in §6 that the epistemic theory provides a philosophical account of causality that is well motivated across a range of sciences.

2 The epistemic theory of causality

In this section, we provide an introduction to the epistemic theory of causality. Although the epistemic theory is a theory of the nature of causality, it is intimately connected to the epistemology of causality, as we shall now describe.

Clearly, we have causal beliefs and make causal claims. These help us to successfully predict, explain and control our world. We shall refer to these predictions, explanations and control inferences that are characteristic of causal claims as 'PECs'.

Our causal beliefs and claims depend on our evidence. In some cases, these beliefs and claims are appropriate, given the evidence, while in others they are not. (Whether causal claims are appropriate given the evidence is a different question to whether they are borne out by further enquiry.) We refer to a theory that says something about which causal claims are appropriate given the available evidence E as a *causal epistemology*.

A variety of causal epistemologies have been put forward in the literature. In the literature on causal cognition, there are accounts that emphasise counterfactual reasoning (Gerstenberg et al., 2022), interventions (Gopnik et al., 2004), temporal cues (Lagnado and Sloman, 2004), mechanistic information (Ahn et al., 1995) and decision making (Sobel and Kushnir, 2006), for example. In science and medicine, there are accounts that emphasise randomised trials (Guyatt et al., 1992; Sackett et al., 1996), quantitative methods (Imai, 2017), qualitative methods (Patton, 1990; Glynn and Ichino, 2015), or mixed methods (Teddlie and Tashakkori, 2009).

This plethora of approaches highlights three points. Firstly, there are clearly many diverse indicators of causality. Second, there is no settled view as to how to capture these indicators in a single causal epistemology. Third, some causal epistemologies may be better than others. In medicine, for example, there has been a transition from an approach based on authority and experience to 'evidence-based' medicine (EBM). This transition has been accompanied by improvements to health outcomes. Proponents of the EBM+ approach to evidence-based medicine have argued that further improvements may be possible, by making additional changes to our causal epistemology (Parkkinen et al., 2018).

This third point raises the possibility that we are progressing towards some optimal causal epistemology—albeit an ideal that may never in fact be reached. What would it take for a causal epistemology to be optimal? Such a theory would need to balance several demands. Putative desiderata include the following. (i) Reliability: a causal epistemology should yield a body of causal claims that underwrites successful PECs. (ii) Strength: it should establish and rule out sufficiently many claims for science to progress efficiently. (iii) Stability: those claims that are established or ruled out should likely remain so in the face of subsequent evidence.[1] (iv) Completeness: it should determine,

[1]Some of these desiderata are general epistemological desiderata, not specific just to causal epistemologies. For example, Stability is a desideratum associated with establishing in

for any putative causal claim, whether the claim is established or ruled out by evidence, or, if neither, what degree of confidence is warranted in the claim. (v) Simplicity: it should not be unnecessarily complicated; e.g., single-case causal claims should be subsumable under generic causal claims as far as possible. (vi) Feasibility: it should be practicable to use the causal epistemology to evaluate causal claims of interest.[2]

According to the epistemic theory of causality, these ingredients are all we need for an adequate theory of causality: *the causal facts are just facts about what is established or ruled out by all optimal causal epistemologies on total evidence.* A is a cause of B just if every optimal causal epistemology would deem the claim that A is a cause of B to be established by an idealised evidence base that consists of all matters of particular fact—past, present and future. A is not a cause of B iff every causal epistemology deems the claim that A is a cause of B to be ruled out by total evidence. Otherwise—if some optimal causal epistemologies deem A to be a cause of B and others not—it is indeterminate whether A is a cause of B.

According to the epistemic theory, then, there is no need to take causality to be 'out there', nor analysable in terms of a single indicator of causality, such as probabilistic dependence, counterfactual connection, mechanistic connection, temporal succession etc. The reality is that all these indicators play a role in our judgements about what causes what, and attempts to give one primacy over the others invariably fail. To understand causality we need to understand the roles these indicators play in fixing our causal beliefs. The facts of causality are facts about rational belief, not facts about some non-epistemological connection between the causal relata.

The epistemic understanding of causality in terms of rational belief is analogous to the Bayesian understanding of probability in terms of rational belief. Thus a parallel can be drawn between epistemic causality and epistemic probability. A probabilistic belief is a kind of belief, not a belief about some kind of non-epistemic probability: it is a relational belief of the form $P_E(A) = x$, which says that rational degree of belief in A on evidence E equals x. Our probabilistic beliefs enable successful predictions, decisions and actions, and the facts about rational probabilistic belief are determined by optimal probabilistic epistemologies, not by a single indicator of probability, such as symmetry, observed frequency, or confirmed theory. Similarly, causal belief is a kind of belief, not a belief about some kind of non-epistemic causal relation: it is a relational belief of the form $C_E(A, B)$, and a body of such beliefs enable characteristic PECs. The analogy between epistemic causality and epistemic probability can be pushed further: Williamson (2021) argues that the two approaches admit analogous norms and yield analogous analyses of the relevant facts.

general, and not specific to establishing causal claims (Williamson, 2022). When establishing a proposition we expect to be able to use it as evidence for other propositions in the long term. It would be almost Moore-paradoxical to say that 'We have established that smoking causes cancer but we expect to retract this claim in the coming year.' While some established claims will inevitably be overturned in the light of new evidence, the likelihood of this happening for any given claim ought to be small.

[2]The difference between Completeness and Feasibility is that the latter requires that the epistemology be one that could be applied in practice, while the former does not.

3 The case for epistemic causality

Having introduced the key tenets of epistemic causality, we shall now briefly sketch four ways of motivating this view of causality: by appeal to the inadequacy of alternative accounts (§3.1), by appeal to parsimony (§3.2), by appeal to a recent view of the epistemology of causality (§3.3), or by appeal to the need to remain neutral about certain questions (§3.4). These arguments will be explored in more detail in subsequent sections.

3.1 The argument from failure

The argument from failure proceeds from the observation that other theories of causality tend to fall to counterexamples, while epistemic causality does not. These considerations favour the latter theory over the former theories.

Standard theories of causality are often classed as difference-making or mechanistic theories. Difference-making theories include the regularity, counterfactual, probabilistic, agency and interventionist theories. All these theories require that a cause should make the appropriate kind of difference to its effects. Unfortunately, one can usually find cases of causation where the cause cannot make a difference, because the effect is already fully determined by other factors (see, e.g., Hall, 2004). One strategy here is to suggest that the cause would made a difference when the other factors are absent, but even this strategy does not work when the set of causes is mutually exclusive and exhaustive, because then the cause of interest cannot be varied independently of the others (Williamson, 2005, §7.3). Mechanistic theories, on the other hand, require that cause and effect should be connected by some appropriate kind of mechanism. These theories face problems in cases where the cause or the effect is an absence of something, since an absence cannot be a part of a mechanism (see, e.g., Hall, 2004; Williamson, 2011, §II.1). The behaviour of a mechanism is supposed to be explained by the arrangement of its (actually present) constituents, such as entities, activities and events.

A common suggestion is to move to pluralism. One approach here is a kind of dualism: the view that some causal claims are claims about difference making while other claims are claims about mechanisms (Hall, 2004). This view falls to counterexamples in which there is neither difference making nor a linking mechanism (Longworth, 2006). Moreover, our use of causal talk stands at odds with pluralism. We do not tend to ask clarifying questions to disambiguate a causal claim, in the way that we might with some probabilistic claim which could be interpreted either as a claim about rational degree of belief or as a claim about frequency. This latter problem also besets more radical kinds of pluralism, such as the inferentialism of Reiss (2012)—see Williamson (2006a, 2013) for further discussion.

Epistemic causality does not succumb to these problems that beset pluralism because it is a monistic theory: although there are multiple indicators of causality (evidential pluralism), there is a single concept of cause (conceptual monism) and a single causal relation (metaphysical monism). Epistemic causality is metaphysically monist because it analyses the causal relation solely in terms of facts about what is established or ruled out by all optimal causal epistemologies on total evidence. There is no other causal relation, according to epistemic causality, and this is evidenced by the decisive objections that face

other accounts of causality. Epistemic causality is conceptually monist because it invokes a single rational-belief concept of cause. Again, no other concept of cause is viable, according to epistemic causality, as is witnessed by objections to other concepts of cause and to pluralism.

Additionally, epistemic causality does not succumb to the problem of counterexamples involving cases of overdetermination or absences. Indeed it is hard to see how one could produce any sort of counterexample to epistemic causality. A counterexample to epistemic causality would require finding a causal fact that the epistemic theory misclassifies. Recall that the epistemic theory deems A to be a cause of B just when every optimal causal epistemology deems that A is a cause of B, on total evidence. It deems A not to be a cause of B iff every optimal causal epistemology deems that A isn't a cause of B. It is indeterminate whether A is a cause of B iff some optimal causal epistemology says it is and some other says it isn't. A counterexample to the epistemic theory would need to show that the theory misclassifies some particular causal claim, and it would need to do this by appeal to some consideration that settles the question of the correct classification. The difficulty is that if this consideration is correct and conclusive then one would expect that it would be validated by each optimal causal epistemology. For instance, if the example is one of causation between absences, and our intuitions about the example are correct, then one would anticipate that every optimal causal epistemology would validate the example. So it is hard to see how any counterexample put forward against the epistemic theory could be conclusive.

Thus, the problems that beset the standard accounts of causality favour the epistemic theory over these rival accounts.

3.2 The argument from parsimony

A second line of argument for the epistemic theory appeals to the idea of parsimony, as follows. Clearly we have causal beliefs and make causal claims and we need to theorise about how best to do this, in order to progress science, medicine, public policy, and our own decision making. So we need causal epistemologies and we need to think about how good they are. Then it is but a small step to epistemic causality: its building blocks are just causal epistemologies and the idea of progress towards an ideal causal epistemology. Moreover, the epistemic theory is successful in yielding the correct judgements of causality, as we saw above when we considered the possibility of counterexamples to epistemic causality. Given that the epistemic theory is itself parsimonious and successful, it would be otiose to posit any additional kind of causality. There is simply no need for a further theory that attempts to analyse causality in terms of one of its indicators, or in terms of a pluralist panoply of indicators. Epistemic causality is all we need.

Let us consider three potential responses to this argument from parsimony.

Firstly, one might ask whether some alternative approach to causality can run the same sort of argument. If we analyse causality in terms of some non-epistemic X, perhaps X-causality would be all we need and it would be otiose to consider epistemic causality. There is an asymmetry here, however. The proponent of epistemic causality is likely to find X-causality to be less parsimonious: the alternative theory of causality requires some appropriate stuff 'out there' to which causality can be ultimately reduced. (This might be pos-

sible worlds, required to underwrite modal difference-making claims, or causal powers, for example.) On the other hand, even an advocate of X-causality should admit that we need to consider causal beliefs and causal epistemologies, whether or not causality is analysable in terms of X.

Second, one might think that even if alternative theories of causality are otiose as analyses of causality, they may yet have some heuristic value. For example, a counterfactual theory of causation can suggest certain strategies for testing causation—strategies which may have led to some improvements to past causal epistemologies. But note that one can admit this heuristic role for alternative theories of causality while taking epistemic causality to be the correct analysis of causality. Thus there is no incompatibility here. On the other hand, the heuristic value of alternative theories is by no means clear cut: identifying causality with counterfactual connection may have hampered more than it has helped, for example. Indeed, proponents of EBM+ might argue that this identification has merely delayed potential improvements to causal epistemology in medicine, by entrenching the position of present-day EBM, which has important limitations.[3]

Third, one might worry that we have a slippery slope: that if we accept that parsimony motivates epistemic causality then we would be forced to accept that parsimony motivates an epistemic theory of everything. But we would not want to be forced to adopt an epistemic theory of tables and chairs, for example. The concern is that this would ultimately be too revisionary.

One can resist an epistemic theory of everything, however, by appealing to simplicity and success. It's much simpler, given the way we think, to construe facts about tables and chairs as facts about things out there rather than as facts about rational table-beliefs and rational chair-beliefs. Moreover, taking tables and chairs to be out there is successful—not prone to counterexamples. Hence there is no need to resort to an epistemic theory of tables and chairs.

Thus, epistemic theories differ from certain global approaches—such as pragmatism, Humean supervenience, and inferentialism—in being appropriate only in those cases in which simpler theories cannot be successfully applied.

3.3 The argument from Evidential Pluralism

A third argument for epistemic causality appeals to one particular causal epistemology, namely Evidential Pluralism. This epistemological theory provides a very general account of the confirmation relationships involved in assessing causal claims—general enough to accommodate many of the indicators of causality introduced in §2. On the one hand, this causal epistemology appears to pose a serious challenge to standard accounts of causality. On the other hand, epistemic causality can accommodate Evidential Pluralism perfectly well. Hence, Evidential Pluralism favours epistemic causality over rival accounts.

Evidential Pluralism provides an account of what one needs to establish in order to establish that A is a cause of B, and an account of what sort of studies

[3]The tendency to conflate causality with some specific indicator of causality, such as counterfactual connection, can be viewed as an instance of what Jaynes (2003) called the 'mind projection fallacy'—the mistake made by construing something that is really epistemic to be a feature of the non-epistemic world. Its refusal to project causal relationships onto the world sets epistemic causality apart from projectivist theories of causality (Beebee, 2015).

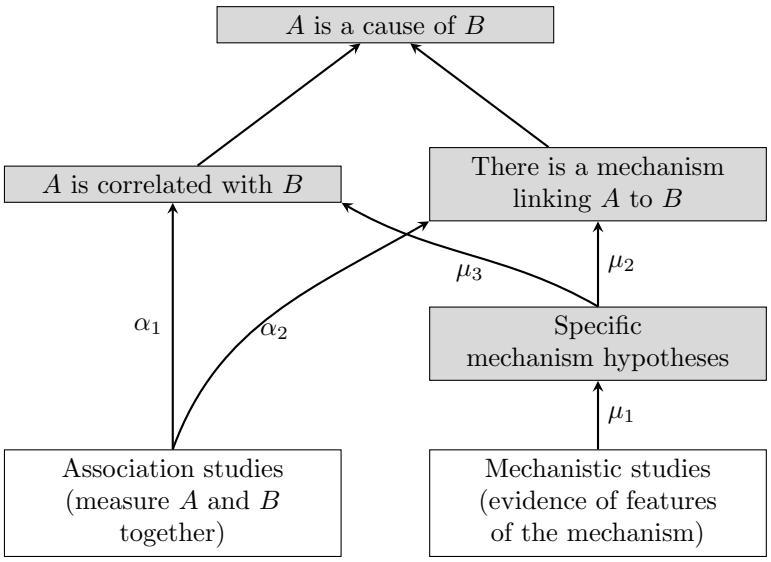

Figure 1: Evidential relationships for establishing a causal claim, according to Evidential Pluralism.

one needs to consider in order to assess causality. Fig. 1 illustrates the main claims of Evidential Pluralism.

Consider first the top part of Fig. 1. Evidential Pluralism is motivated by the platitude that there is more to causation than correlation. What else is needed to establish a causal claim, other than evidence of a correlation conditional on potential confounders? Arguably, a correlation between A and B is attributable to causation just where there is some complex of mechanisms linking A to B according to which instances of A are partly responsible for instances of B. This motivates the thesis that establishing causality requires establishing both the existence of a correlation and the existence of an appropriate mechanistic connection (Russo and Williamson, 2007).

Moving down Fig. 1, we turn to the question of how to establish correlation and mechanism. The usual way to establish correlation is to perform a study that repeatedly measures A and B to test for a correlation (confirmation channel α_1). Certain kinds of these association studies can also indirectly confirm the existence of a linking mechanism (α_2). In particular, high quality randomised controlled trials can confirm the presence of a linking mechanism, by making it less likely that an observed correlation is attributable to confounding. But there is a more direct way to establish the existence of a suitable mechanism: hypothesise key features of the mechanism and perform studies that test for the presence of these features (μ_1 and μ_2). In certain cases—especially if the details of the mechanism are well established and the mechanism itself is not too complex—this mechanistic evidence can also make the existence of a correlation more plausible (μ_3).

Evidential Pluralism poses a challenge for standard accounts of causality (Russo and Williamson, 2007). If a difference-making theory of causality were correct, it should be sufficient to establish the appropriate sort of correlation

in order to establish causation: there should be no need to go on to establish the existence of a mechanism. A similar point holds for any mechanistic account of causality: if such a theory were correct, there should be no need to establish correlation once the appropriate sort of mechanism were established. Dualist accounts face the same problem. For a dualist account, some claims are difference-making claims, while others are mechanistic claims. For those that are difference-making, it should be sufficient to establish correlation. For those that are mechanistic, it should be sufficient to establish mechanism. Standard theories, then, need to either refute Evidential Pluralism or provide an account of how to accommodate Evidential Pluralism. To date, this challenge has not been adequately met.[4]

On the other hand, it is straightforward for the epistemic theory of causality to accommodate Evidential Pluralism. Indeed, a key feature of the epistemic theory is its flexibility to cope with multiple indicators of causality. If any ideal causal epistemology validates Evidential Pluralism, then, by construction, epistemic causality also validates Evidential Pluralism. Thus Evidential Pluralism favours epistemic causality over a range of standard rivals.

3.4 The argument from neutrality

Certain contentious positions in science are allied with either a difference-making account of causality or a mechanistic account. In order not to prejudge these contentious questions, it is important not to side with some such account. Epistemic causality does not presuppose that all causal relationships are difference-making relationships, nor that all causal relationships are mechanistic relationships. Thus the need for neutrality can favour the epistemic theory over standard alternative theories of causality.

For example, Taylor (2021) argues that epistemic causality is needed to provide a unified conception of causal explanation in cognitive science, which admits the possibility of both mechanistic and non-mechanistic explanations of cognition. The idea is that epistemic causality is required so as not to prejudge the question of whether all causal explanations in cognitive science are mechanistic: a mechanistic theory of causality would say yes, while a difference-making or dualist approach would say no. Since this question has not been settled in cognitive science, we are not in a position to ascertain whether ideal causal epistemologies will have a role for non-mechanistic causal explanations. Hence, the epistemic theory remains neutral on this question. Arguably, one needs a theory of causality that remains neutral on this question, precisely because the question hasn't been settled. This favours the epistemic theory over standard rivals.

We will consider another instance of this argument, as well as an instance of the argument from Evidential Pluralism, when we consider the cognitive sciences in §5. Meanwhile, in the next section, we will see how some of these arguments play out in the social sciences.

[4]Though see Weber (2009) for an account of how Giere's probabilistic theory of causality might meet this challenge.

4 Epistemic causality in the social sciences

In this section, we argue that the epistemic theory provides an account of causality that is well suited to the social sciences. We appeal here to the arguments from failure, parsimony, and Evidential Pluralism.

There are two standard approaches to causality in the social sciences: monism and pluralism (cf. Gerring, 2005; Haggard and Kaufman, 2016). First, in §4.1, we shall argue that the epistemic theory is preferable to the standard monistic approach by appeal to the argument from failure. Next, in §4.2, we shall show that the epistemic theory is preferable to the standard pluralist approach by appeal to the arguments from parsimony. Finally, in §4.3, we shall argue for the epistemic theory by appeal to the argument from Evidential Pluralism.

4.1 Causal monism in the social sciences and the argument from failure

The standard monistic approach in the social sciences typically understands causality in a difference-making way (e.g., Granger, 1969, 1980; King et al., 1994; Morgan and Winship, 2015).[5] For example, King et al. (1994, pp. 81–82) define causality in terms of 'the difference between the systematic component of observations made when the explanatory variable takes one value and the systematic component of comparable observations when the explanatory variable takes on another value'.[6] While some authors adopt a probabilistic account of difference making, others appeal to counterfactuals (e.g., Cook and Campbell, 1979; Holland, 1986): they take causes to be factors that are manipulatable in experiments.

These difference-making approaches face serious objections. Typical objections to the probabilistic account stem from the truism that correlation is not causation. For example, it has been shown that there is a statistically significant correlation between unemployment rates and property crime in the 1990s in the US (Gould et al., 2002; Machin and Meghir, 2004). A typical estimate would be that a one percentage point increase in the unemployment rate is associated with a one percent increase in property crime. Based on these estimates, the observed two percentage point decline in the US unemployment rate between 1991 and 2001 has been taken to cause the estimated two percent decline in property crime. However, as Levitt (2004) indicates, the correlation between unemployment and crime rates in this case is historically contingent. As a counterexample, the 1960s is a decade of strong economic growth with a sharp increase in crime rates. Instead, Donohue and Levitt (2001) argue that the drop in crime rates in the 1990s in the US was caused by the legalisation of abortion in the 1970s. Thus, as has been well recognised, the probabilistic variant of the monistic approach is vulnerable to the problem that establishing correlation is not sufficient for establishing causation.

[5] Social scientists who embrace difference-making theories do not necessarily have a strong views about the ontology of causation. For example, those who advocate a probabilistic account of causation do not necessarily hold views about what constitutes this probabilistic dependence.

[6] Although King et al. (1994, pp. 86–87) recognise the significance of mechanisms in causal analysis, they contend that the concept of causality is essentially difference-making rather than mechanistic.

The counterfactual account can also be challenged. There is a worry concerning the assumption of the manipulability of causal variables. According to some counterfactual accounts, all causal factors must be experimentally manipulable.[7] In other words, social scientists should not assert any causal claim about non-manipulable variables. As Goldthorpe (2001, p. 6) illustrates, 'one could discuss the association that exists between sex or race, on the one hand, and say, educational attainment, on the other. But it would be no more meaningful to speak of sex or race as being causes of such attainment than it would be to make statements about what level of education Ms M would have achieved had she been a man or Mr N had he been a woman.' If that is right, the scope of causal claims in the social sciences is much more limited than is commonly thought. In addition, there is a concern about the distinctive nature of the response of the units in experiments in the social sciences. The counterfactual account of causation was originally introduced to the social sciences, especially sociology, following its application to medical and agricultural science (Cook and Campbell, 1979). However, this approach to causality in applied natural science cannot be straightforwardly transported to a sociological context. In principle, the counterfactual account allows conceptual space for human action only in the role of experimenter or intervener. For example, in an experiment to test a fertilizer, the experimental set-up is the only source of intervention. Once the experiment is carried out, all else has to follow in the manner of plants responding to the fertilizer. But in the social sciences, the response of the units in experiments cannot be simply assumed to have the same nature as that of the units in experiments in applied natural and agricultural sciences. Consider a case of the introduction of positive discrimination in education, with the aim of reducing class or ethnic differentials in achievement. It is likely that members of those classes or ethnic groups whose children would not benefit and who might lose their competitive advantage in schools could respond in order to preserve the advantage. In this case, one crucial requirement of experimental design would be breached: the response of a unit should not be influenced by whether other units are treated or not.

The epistemic theory does not succumb to these problems, because it does not analyse the causal relation in terms of difference making and it does not impose strict conditions on what can count as causal relata. Thus, the counterexamples to the above monist accounts cannot pose a challenge to epistemic causality. These problems can be viewed as illustrating the argument for epistemic causality from the failure of standard alternatives.

Probabilistic and counterfactual accounts of causation are not the only monistic approaches that are relevant to the social sciences. Political science and sociology might seem to presuppose a mechanistic account of causality (Beach and Pedersen, 2013; Hedström and Ylikoski, 2010). For example, the use of process tracing, case studies and certain small-N studies might seem to establish causation just by establishing mechanism. Just as causation is not correlation, however, so causation is not easily reducible to mechanistic connection. Indeed, as Shan and Williamson (2023) argue, while specifying a mechanism can provide a narrative explanation, it falls short of establishing a

[7] This applies to counterfactual accounts of causation (e.g., Rubin, 1974; Holland, 1986) which originated in the agricultural sciences, but arguably not to Lewis's counterfactual theory (Lewis, 1973) nor Woodward's manipulation theory (Woodward, 2003).

causal claim. For example, political scientists might detail the social actors and their activities of a mechanism from the assassination of Archduke Ferdinand to the outbreak of World War I, but this does not establish that the assassination of Archduke Ferdinand caused the outbreak of World War I, because the war might have happened anyway.[8]

It is evident that this problem does not undermine the epistemic theory, because the epistemic theory does not define causality in terms of mechanism. The epistemic theory is thus immune to key problems that beset monistic accounts of causality because it does not conflate causation with any particular indicator of causation.

4.2 Causal pluralism in the social sciences and the arguments from parsimony

Pluralists maintain that there are multiple concepts of causality in the social sciences (e.g., Reiss, 2009; Goertz and Mahoney, 2012; Maziarz, 2020; Rohlfing and Zuber, 2021). For example, Goertz and Mahoney (2012) argue that there are two concepts of causality in the social sciences that underlie the two key approaches to social scientific practice, namely the quantitative and qualitative approaches. Goertz and Mahoney's argument rests on their own interpretation of Hume's theory of causality. Hume's definition of cause is as follows:

> [W]e may define a cause to be an object, followed by another, and where all the objects, similar to the first, are followed by objects similar to the second. Or, in other words, where, if the first object had not been, the second never had existed. (Hume, 1748, §7.2.29.)

Traditionally, the first part of this definition is regarded as the original formulation of the regularity theory of causality, while the second part is viewed as a precursor to the counterfactual theory of causality. Goertz and Mahoney (2012, p. 76) follow this received view by calling the first part 'the constant conjunction definition' and second part 'the counterfactual definition'. However, Goertz and Mahoney argue that these two definitions correspond to the respective views of causation of the quantitative and qualitative traditions in the social sciences:

> Hume's famous quotation contains two definitions of causation. Definition 1 suggests a constant conjunction between cause and effect, such that effects always follow causes. This definition assumes many cases and has affinities with quantitative views on causation. Definition 2 suggests a counterfactual view of causation, in which the absence of a cause leads to the absence of an outcome. This definition is built around a single case and has important linkages to qualitative views of causation. (Goertz and Mahoney, 2012, p. 81.)

For Goertz and Mahoney, the constant conjunction definition (or definition 1) fits quantitative methods, which presume a statistical approach to establish-

[8]For more in-depth discussion of the distinction between narrative and causal explanation, see Shan and Williamson (2023, pp. 126–127).

ing causal claims, while the counterfactual definition (or definition 2) fits the methods of the qualitative approach better.[9]

As Gerring (2005) indicates, the pluralist approach overstates the ontological, epistemological, and conceptual differences between causal analyses in the social sciences. Let us illustrate this problem with a famous example of sociological research: the study of socioeconomic status and health status (House et al., 1994; Link and Phelan, 1995; Adler and Newman, 2002; Pampel et al., 2010; Phelan et al., 2010). It has been shown that there is a strong association between socioeconomic status and health status. For example, lower socioeconomic status is associated with the 14 major causes of death in the International Classification of Diseases (Illsley and Mullen, 1985). In addition, lower socioeconomic status is shown to be associated with lower life expectancy, higher overall mortality rates, and higher rates of infant and perinatal mortality (Dutton, 1986; Adler et al., 1994; Bosworth, 2018). However, it is debatable whether socioeconomic status is a cause of health status. Sceptics typically argue that socioeconomic status is a placeholder variable for real causes of diseases that have not yet been identified. Thus Rothman (1986, p. 90) suggests that socioeconomic status is 'a correlate of many causes of diseases'.

Even for some social scientists who argue for the causal relationship between socioeconomic status and health, a strong and pervasive association between socioeconomic status and health merely provides 'a description of the social patterning of disease' (Link and Phelan, 1995, p. 82). It is widely accepted that in order to establish the causal claim that socioeconomic status is a cause of disease, one has to establish the existence of some mechanism as well as a correlation (House et al., 1994; Phelan et al., 2004). As Link and Phelan (1995, p. 82) suggest, it is necessary to identify 'the direction of causation between social conditions and health and the mechanisms that explain observed associations' for the purpose of 'establishing a causal role for social factors'.

With their collaborators, Link and Phelan have identified a variety of mechanisms linking socioeconomic status to health status (Link and Phelan, 1995; Phelan et al., 2004, 2010). It is shown that people of higher socioeconomic status possess a wide range of resources, including money, knowledge, power and beneficial social connections, which shape health-enhancing behaviours (such as getting flu jabs, eating fruits and vegetables, and exercising regularly) and access to broad contexts that are associated with risk and protective factors of health. For example, those who have lower status jobs more commonly have 'job strain' (i.e., a combination of high job demands and low decision latitude), which is associated with coronary heart disease (Schnall et al., 1990); people with lower socioeconomic status are more likely to smoke and be overweight, which lead to various health problems (Lantz et al., 1998); and those with lower socioeconomic status experience greater residential crowding and noise, which is linked to poorer long-term memory and reading deficits (Evans and Saegert, 2000).

Moreover, Phelan and Link argue that although there are various mechanisms linking socioeconomic status and health status, no individual mechanism is so dominant that it alone is responsible for the bulk of the observed asso-

[9] It should be noted that Goertz and Mahoney (2012, pp. 81–82) also indicate that some qualitative researchers, especially those who use qualitative comparative analysis, 'may gravitate' towards definition 1.

ciation. In other words, there may be different mechanisms underlying the association between socioeconomic status and health status over time. As Lutfey and Freese summarise, 'the association persists even while the relative influence of various proximate mechanisms changes' (Lutfey and Freese, 2005, p. 1328). It is in this sense that socioeconomic status is a 'fundamental cause' of health status, which is the key idea of the so-called theory of fundamental causes (Link and Phelan, 1995; Phelan et al., 2010).

It is evident that this case is difficult to characterise in terms of the pluralist approach. Indeed, Phelan, Link, and their associates do not take their study to establish two types of causation or two distinct causal claims. They contend that their study successfully identifies socioeconomic status as a 'fundamental cause' of health status (Link and Phelan, 1995, p. 80), which is not easily understood in a pluralist sense.

Suppose there are two different concepts of causality in the social sciences, say, difference-making causality ('$\text{causes}_{\text{DM}}$') and mechanistic causality ('$\text{causes}_{\text{Mech}}$'). If so, '$A$ $\text{causes}_{\text{DM}}$ B' says something different to 'A $\text{causes}_{\text{Mech}}$ B'. However, this is difficult to square with the study of socioeconomic status and health status. Link and Phelan (1995) have shown that there is a strong correlation between socioeconomic status and health status and there are some established mechanisms linking socioeconomic status to health status. What can we conclude from this? Is socioeconomic status a cause of health status? If so, in what sense?

1. Does the socioeconomic status cause_{DM} health status?
2. Does socioeconomic status $\text{cause}_{\text{Mech}}$ health status?
3. Or, do Link and Phelan's studies suggest a new concept of causality?

It seems that the pluralist approach to causality in the social sciences leads to greater confusion. What is worse, the pluralist approach may lead to a problem of incommensurability. As Gerring argues,

> If causation means different things to different people then, by definition, causal arguments cannot meet. If A says that X_1 caused Y and B retorts that it was, in fact, X_2 or that Y is not a proper outcome for causal investigation, and they claim to be basing their arguments on different understanding of causation, then these perspectives cannot be resolved; they are incommensurable. (Gerring, 2005, p. 165.)

Therefore, as Gerring (2005, p. 190) argues, 'pluralistic views are either unconvincing or, to the extent that they are true, unfortunate. We need a single framework within which to understand causal relationships in the social sciences.'

The epistemic theory provides the required simple and unified framework for understanding causal relationships in the social sciences. As we argued in §3.2, the epistemic theory obviates the need for a pluralist approach that invokes multiple concepts of cause. Even though epistemic causality invokes a single concept of cause, it can yield the correct causal judgements. This can be viewed as an instance of the argument from parsimony.

4.3 Causal pluralism in the social sciences and the argument from Evidential Pluralism

Phelan and Link's study fits epistemic causality well because it fits Evidential Pluralism: correlation and mechanism are established as a means to establish causation. This study is not an isolated case. In causal enquiry, it is not unusual that social scientists look for both types of evidence to support their causal claims, instead of focusing on one or the other. Other famous examples include Donohue and Levitt's study of legalised abortion and crimes and Weinstein's study of rebellion and abortion (Shan and Williamson, 2021). As argued in §3.3, the fact that causal enquiry accords with Evidential Pluralism favours epistemic causality over rival accounts such as a difference-making monistic account or causal pluralism.

Phelan and Link's study can thus be used to exemplify both the argument from Evidential Pluralism and the argument from parsimony. Both the monistic approach and the pluralist approach have difficulties in accounting for the causal analysis of socioeconomic status and disease, while epistemic causality does not. Epistemic causality provides a more parsimonious understanding of Phelan and Link's study, and this study also accords well with Evidential Pluralism, which again favours epistemic causality.

Some might argue that in certain cases social scientists do indeed employ different methods to establish different causal claims. For example, in political science, some tend to use process tracing alone to make causal inferences, while others employ statistical techniques to establish causal claims. These methods are so different that one might reasonably infer that causal claims established using these methods appeal to different concepts of cause. If this is the case, the pluralist approach provides a better explanation than the epistemic theory (at least in some cases).

This argument is basically an inference from methodological diversity to causal pluralism (Maziarz, 2021).

> Methodological diversity: there are different approaches to establishing causal claims in the social sciences.
>
> Causal pluralism: there are different concepts of causality in the social sciences.

While the methodological diversity thesis seems to be true, causal pluralism is much more doubtful. Causal pluralism is certainly not required to explain methodological pluralism in the social sciences—epistemic causality provides an alternative account. The different methodological approaches to establishing causal claims can be understood as different ways to obtain evidence for monistic causal claims. Consider Weinstein' study of rebellion and violence. Weinstein (2007) argues for the causal claim that the initial conditions that rebel leaders encounter cause their strategy of violence. In order to support this causal claim, 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 the case studies of rebel groups in Mozambique, Peru, and Uganda. In short, Weinstein uses both statistical techniques and ethnographic methods to establish a causal claim.

As argued by Shan and Williamson (2023), statistical techniques and ethnographic methods are not used to establish different types of causal claims. Rather they are used to obtain different objects of evidence to justify a causal belief. Therefore, methodological diversity does not support causal pluralism. Rather, it reflects Evidential Pluralism.

In addition, Shan and Williamson (2021) argue that even in the cases where political scientists are using process tracing alone to establish causal claims, evidence of correlation is assumed, though often implicitly. As Gerring (2005, p. 166) indicates, 'some correlational-style analyses slight the explicit discussion of causal mechanisms but this is usually because the author considers the causal mechanism to be clear and hence not worthy of explicit interrogation. Similarly, a mechanistic argument without any appeal to covariational patterns between X and Y does not make any sense. The existence of a causal mechanism presumes a pattern of association between a structural X and an ultimate Y.' Thus the use of process tracing methods does not imply that the concept of causality is mechanistic in nature. Nor does the use of statistical methods suggest that the concept of causality is fundamentally correlational. As we have argued, they can be understood as attempts to obtain different objects of evidence. As Crasnow (2011, p. 47) notes, 'pluralism about methodology need not commit us to a conceptual pluralism about causes'.

In sum, the epistemic theory provides an account of causality that fits well with causal enquiry in the social sciences. First, it does not fall to counterexamples that beset monistic accounts. Second, the epistemic theory provides a simpler and more unified account of causality than causal pluralism. Third, it fits very well with Evidential Pluralism, which provides a more tenable epistemological account of causality than causal pluralism.

5 Epistemic causality in the cognitive sciences

Following Thagard (2005), we can define cognitive science as the interdisciplinary study of mind, embracing philosophy, psychology, artificial intelligence, neuroscience, linguistics, and anthropology. The explanatory results of cognitive science can then be understood as "theoretical and experimental convergence on conclusions about the nature of mind," which—as with any science—must be framed in terms of empirically supported explanations and predictions.

A central task of cognitive science is to develop causal explanations (Cummins, 2000; Kaplan and Craver, 2011; Taylor, 2021). In some instances, cognitive theorists—such as psychologists—try to explain behaviour. For example, Piccinini and Craver (2011, 283) argue that:

> When psychologists explain behavior, the explanations typically make reference to causes that precede the behavior and make a difference to whether and how it occurs. For instance, they explain that Anna ducked because she saw a looming ball.

But cognitive theorists do not only try to explain behaviour; they also aim to develop causal explanations of the system—e.g. the mind/brain—that is responsible for cognitive phenomena. In this vein, cognitive theorists have formulated a range of causal explanations of cognitive competences; for instance, causal explanations of categorisation (cf. Davidoff, 2001; Harnad, 2017; Lin and

Murphy, 1997; Taylor and Sutton, 2021), perception (Chater and Vitányi, 2003; Sims, 2018; Tanrıkulu et al., 2021), and memory (Baddeley, 1992; Michaelian and Sutton, 2013; Morrison and Chein, 2011).

Despite the focus on causal explanations in cognitive science, little attention has been paid to the epistemic theory of causality. Instead, the focus has been predominately on mechanistic (cf. Kaplan and Craver, 2011; Piccinini and Craver, 2011) and difference-making (e.g., interventionist) theories of causality (Meyer, 2020). We think that this is a mistake, because the epistemic theory of causality provides a very natural account of causality in cognitive science. Endorsing the epistemic theory of causality can help in at least two ways. First, to accommodate the epistemology of causality in cognitive science, which can be seen to conform to Evidential Pluralism (§5.1). Second, to allow us to remain neutral about long-standing tensions in cognitive science concerning how to individuate causally efficacious mental states (§5.2).

5.1 Evidential Pluralism in Cognitive Science

We suggested in §2.3 that the epistemic theory of causality best accommodates Evidential Pluralism. Our claim here is that many working cognitive scientists conform to Evidential Pluralism and, as a result, the theory that best makes sense of their attempts to establish causal claims is the epistemic theory of causality.

As a reminder, Evidential Pluralism is the view that:

> In order to establish that A is a cause of B one normally needs to establish two things. First, that A and B are suitably correlated—typically, that A and B are probabilistically dependent, conditional on B's other known causes. Second, that there is some underlying mechanism linking A and B that can account for the difference that A makes to B.

Evidential Pluralism has already found good support in the health and social sciences, but open questions remain about its applicability elsewhere. Here, we argue that Evidential Pluralism is applicable to the cognitive sciences. To make this case, we consider two examples from cognitive neuroscience and developmental psychology respectively: Dehaene's (2009) theory of reading and discussions of "theory of mind."

Before turning to these examples, it is important to make one point explicit; namely, that in this subsection, we will be primarily concerned with demonstrating that Evidential Pluralism is, in fact, the causal epistemology that best describes cognitive scientific practice and, hence, that Evidential Pluralism is the causal epistemology that should be endorsed in (at least) this context. It follows from this that we have further support for the claim that the epistemic theory of causality provides the best account of causality in cognitive science, because we have already said (see §3.3 above) that the epistemic theory of causality accommodates Evidential Pluralism and that other theories of causality will either refute Evidential Pluralism or have yet to provide an account of how they accommodate Evidential Pluralism.

Our first example is taken from Dehaene (2009, 5) who argues that "the brain contains fixed circuitry exquisitely attuned to reading" and that the

functional activity of this cortical area is causally responsible for our capacity to recognise words and letters. The area in question is located in the left ventral occipito-temporal junction and is now commonly labelled with a functional designation that Dehaene himself coined: the visual word form area (VWFA). Dehaene's idea is that the function of VWFA is causally responsible for certain behaviours (namely, letter/word recognition and reading), because patterns of activity in VWFA play a causal role in the cognitive process that enables the organism to read by acting as a signal that informs the activities of downstream neural mechanisms.

Thus, Dehaene makes a causal claim: that patterns of activity in the VWFA (in response to certain environmental parameters) cause cognitive processes that enable organisms to recognise letters/words and, ultimately, to read. The open question, however, is how this causal claim is established. And it is here that Evidential Pluralism seems to be the causal epistemology at work. The reason is that Dehaene only feels able to put forward the aforementioned causal claim after establishing *both* mechanism and correlation.

In this case, the evidence of correlation is plentiful. For example, there is evidence that in normal literate subjects, VWFA is differentially responsive to written, but not spoken words (Dehaene and Cohen, 2007); that in blind subjects, the region is differentially responsive to words presented in Braille, but not to tactile control stimuli (Reich et al., 2011); and that lesions to VWFA appear to result in pure alexia, a condition in which formerly literate subjects cannot understand written words, despite being able to understand and produce verbal speech at roughly normal levels of competency (Gaillard et al., 2006). In each of these examples, we find that the functioning of VWFA (or not) makes a difference to our capacity to recognise letters/words and to read. This is exactly what Evidential Pluralism takes to be evidence of correlation.

But Dehaene also ensures that there is adequate evidence of mechanisms; specifically, evidence of the mechanisms for reading that link the functional activities of VWFA to the functional activities of other areas of the brain (see Figure 2). In particular, Dehaene (2009, 75) appeals to evidence of mechanisms from the mechanistic studies of Dehaene et al. (2002) in support of the following specific mechanism hypothesis:

> The left occipitotemporal "letterbox" [i.e., VWFA] identifies the visual form of letter strings. It then distributes this invariant visual information to numerous regions, spread over the left hemisphere, that encode word meaning, sound pattern, and articulation. [...] Learning to read thus consists in developing an efficient interconnection between visual areas and language areas. All connections are bidirectional.

This evidence of the organisation and activities of different regions of the brain (e.g., posterior parietal region, occipital regions) linking VWFA to our capacity to read goes beyond mere association and has been supported by a number of mechanistic studies. In fact, in their conclusion of one such mechanistic study, Cohen et al. (2002, 1066) argue that:

> The fact that the location of the VWFA is highly reproducible across subjects suggests that some initial properties intrinsic to this

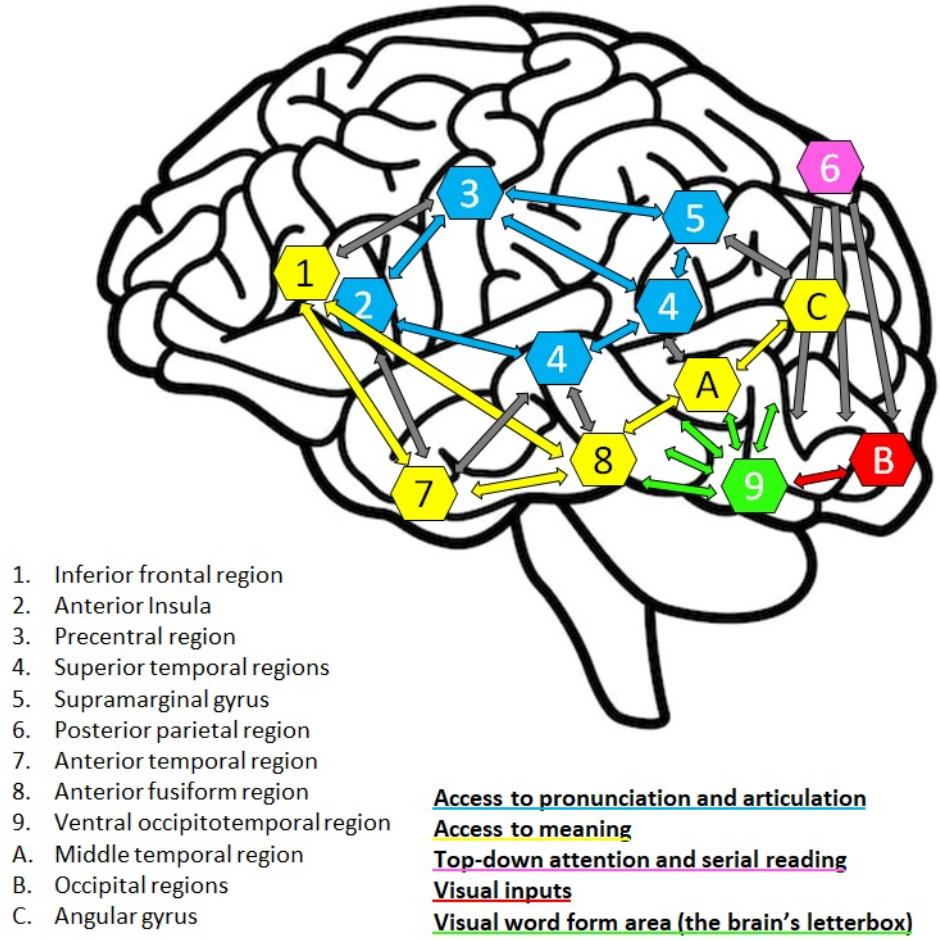

Figure 2: A modern vision of the cortical networks for reading. Adapted from Dehaene (2009, 84, fig. 2.2).

> region and to its pattern of connectivity are the cause of its subsequent specialization for reading.

We find, therefore, that Evidential Pluralism accords with the causal enquiries of working cognitive neuroscientists like Dehaene. But Evidential Pluralism also accords with the causal epistemology of another area of cognitive science: developmental psychology. In particular, the causal epistemology of Evidential Pluralism is at play in discussions about causal claims related to the "theory of mind" (ToM) hypothesis.

According to the earliest account of ToM in Premack and Woodruff's (1978) study of the mind of chimpanzees:

> In saying that an individual has a theory of mind, we mean that the individual imputes mental states to himself and to others (either to

> conspecifics or to other species as well) (Premack and Woodruff, 1978, 515).

ToM has been postulated to play a causal role in belief formation and action, which has been studied via so-called "false-belief tasks," such as the Sally-Anne task described below:

> Children are told a story in which Sally places a marble in a basket. Anne then moves the marble to a box while Sally is absent. Children are asked where Sally will look for her marble when she returns (action prediction) or simply where Sally thinks her marble is (belief). Normally developing children as young as 4 years typically pass such tasks, whereas children younger than 4 and much older children with autism typically fail (Leslie et al., 2004, 515).

Thus, one (of many) causal claims involving ToM is that processes underlying ToM cause children to be successful at false-belief reasoning or that processes underlying ToM cause children to pass false-belief tasks such as the one described above.

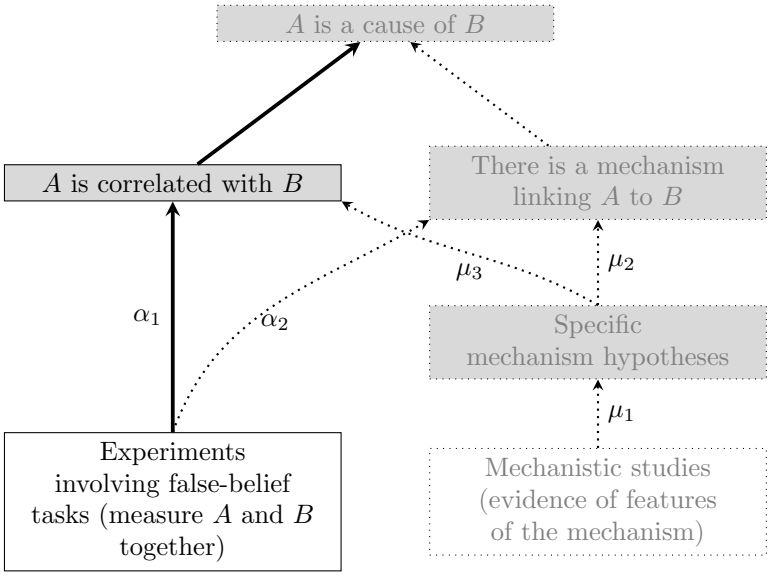

Figure 3: Evidence from experiments of a correlation between being a certain age (and, hence, possessing ToM) and success at false-belief tasks.

To date, a large number of experiments involving false-belief tasks and their variants have been conducted and these seem to have shown that ToM emerges as causally efficacious mental states at around the age of 4. The problem, however, is that such experiments only provide evidence of a correlation between being a certain age (and, hence, possessing ToM) and success at false-belief tasks. This is the case because the results of these experiments are potentially subject to confounders that undermine the external and internal validity of the results.

For example, some have argued that we must be cautious when generalising the results of experiments involving false-belief tasks to other domains, because such experiments have focused almost exclusively on subjects from WEIRD (Western, Educated, Industrialized, Rich, and Democratic) societies (Heinrich et al., 2010). Moreover, some have argued that the results of experiments involving false-belief tasks do not adequately test for the causal role of ToM, because standard false-belief tasks are "verbally based" and so we cannot rule out the possibility that the results are biased by factors related to children's linguistic abilities (Southgate et al., 2007; Onishi and Baillargeon, 2005; Surian et al., 2007). The point, then, is that the experimental results of false-belief only provide strong evidence along channel α_1 of Figure 3 above (see Figure 3 where A is being a certain age (and, hence, possessing ToM) and B is success at false-belief tasks).

Now, if correlation were sufficient for causality, the evidence of correlation obtained via experiments involving false-belief tasks should be enough to establish the causal claim that being a certain age (and, hence, possessing ToM) (A) causes success at false-belief tasks (B). But this is certainly not how practising developmental psychologists have seen the matter. Instead, they have been reluctant to assert that the causal claim has been established until they have acquired evidence of mechanisms. This is clear if we take a brief look at the recent history of research related to theory of mind mechanisms (ToMM).

According to Frith and Happé (1999, 82):

> The cognitive processes which underlie the development of Theory of Mind (ToM) are still a matter of debate. The field can be divided into those who favour a more general explanation for ToM (e.g. simulation, general theory building), and those who argue for the necessity of a dedicated cognitive mechanism (for debate see, for example, Goldman (1993); Gopnik (1993); and chapters in Carruthers and Smith (1996)).

Both camps accept the results of experiments involving false-belief tasks, but disagree about the nature of the cognitive mechanisms underlying ToM that link it to success at false-belief tasks. For some, the mechanisms underlying ToM are "domain general" in the sense that they are not specifically and uniquely tied to our ToM-based ability to attribute mental states to others and predict their behaviour accordingly. For others, however, there is a dedicated—and, perhaps, innately specified—cognitive mechanism underlying ToM.

As an example, consider the claim by Leslie (1987) that "basic representational structures for a theory of mind are put in place by the emergence of the decoupler mechanism." This view—which falls into the domain-specific camp—holds that the origins of ToM can be found in various "decoupling mechanisms," which underpin the ability to represent mental states 'decoupled' from reality. For example, the decoupling mechanism for pretense (see Figure 4), which, according to Leslie (1987, 419–420), operates as follows:

> First, there are the perceptual processes whose job is to feed representations of the current situation to the central processes. Second, there is the set of processes labeled central cognitive systems. These include structures corresponding to perceived situation, memory

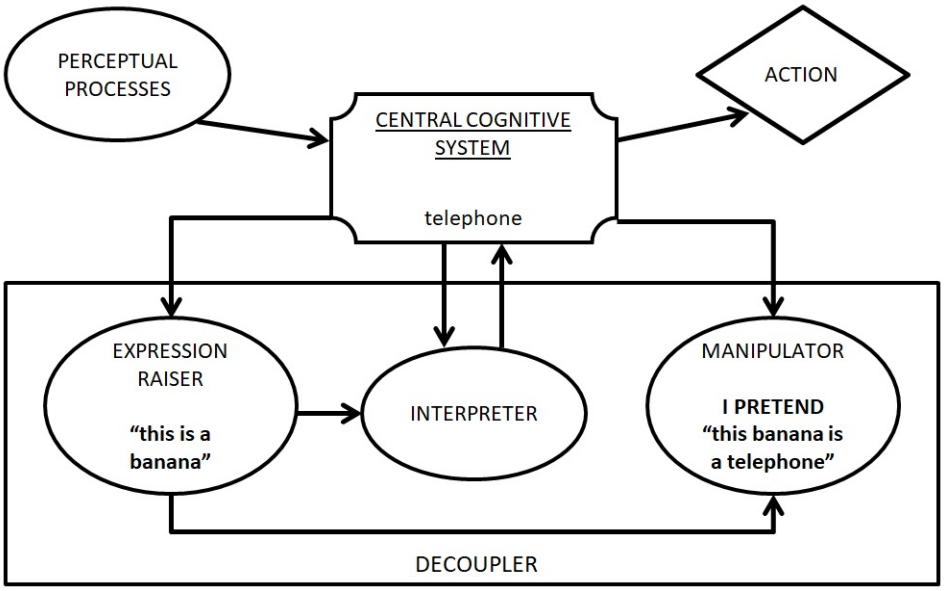

Figure 4: The decoupler model of pretense. Adapted from Leslie (1987, 419, fig. 2).

> systems (including, for example, general knowledge), systems for planning action, and so on.[...]
>
> The expression raiser's job is to copy primary representations from the central systems. It raises copies into the opaque context of the decoupling marks. The copy of the primary expression is thus removed from its normal input-output relations and from the normal computational consequences it would otherwise have. It will now form the nucleus of a metarepresentation [...]
>
> The manipulator's job is to transform decoupled expressions by integrating (primary) information from memory within the decoupling marks or by applying inference rules from memory. [...]
>
> The interpreter can access primary representations in central systems. It performs anchoring functions and relates decoupled expressions to the current perceptual representation. It can access inference rules and other information for passing to the manipulator in a further cycle.

For Leslie, it is the emergence of metarepresentation through the decoupling mechanism that accounts for abilities associates with ToMM and, ultimately, for success at false-belief tasks.

Leslie's account counts as a mechanism hypothesis in our sense and this hypothesis has now found support from several functional neuroimaging studies of ToMM, which purport to identify the neural substrates of "mentalizing" abilities associated with ToM. According to a review by Gallagher and Frith (2003, 78), "the findings of these studies indicate that this ability is mediated

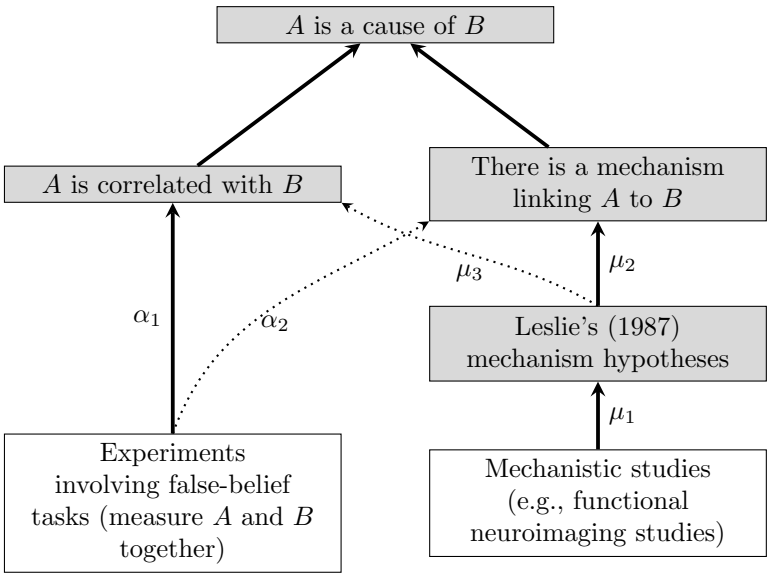

Figure 5: Establishing that being a certain age (and, hence, possessing ToM) causes success at false-belief tasks.

by a highly circumscribed region of the brain, the anterior paracingulate cortex (approximately corresponding to Brodmann area (BA) 9/32)." Moreover, evidence has now been gathered to link Leslie's ToMM to other relevant mechanisms of inhibitory selection, where certain default beliefs are inhibited so that other beliefs can be selected.

Thus, those in favour of the view that there are dedicated and, perhaps, domain-specific cognitive mechanisms underlying ToM will likely assume that we have evidence of a mechanism linking being a certain age (and, hence, possessing ToM) (A) and success at false-belief tasks (B) (see Figure 5). And this counts as evidence of mechanisms along channels μ_1 and μ_2, which combines with evidence of correlation along channel α_1 and so allows us to establish that A is a cause of B. This is exactly why Leslie (1987, 423) asserts that her account "builds a powerful causal story" by building upon the evidence garnered from experiments involving false-belief tasks.

That said, open questions remain about whether it is right to think of ToMM in the dedicated and domain-specific way Leslie favours. For those who are sceptical of this approach, it seems unlikely that the evidence of mechanisms cited above would be seen as definitive. Instead, advocates of domain-general views of ToMM typically call for further research to move beyond the state of affairs represented by Figure 3 above. For example, research to study "tasks based on a refined analysis of specific component processes of mentalizing [...], rather than fixating on the umbrella term ToM" (Schurz and Perner, 2015, 1610).

This demonstrates that for advocates of both domain-general and domain-specific ToMM, the evidence of mechanisms can never be enough to establish causal claims on its own. Every theorist in this area refers back to experimental

results delivering evidence of correlation, because without referring to these results there is nothing to determine whether there is a net correlation. This follows because we cannot undertake, for instance, functional neuroimaging without first specifying the task that participants are to be doing while we scan their brains. And the only way to identify the relevant tasks in this domain is to first obtain evidence of a correlation between some putative cause A—e.g., being a certain age (and, hence, possessing ToM)—and some effect B—e.g., success at false-belief tasks.

Thus, the lesson is clear: developmental psychologists studying ToM are not willing to assert that a causal claim has been established until they have evidence of *both* correlation and mechanism. Neither will be sufficient in isolation. This accords perfectly with Evidential Pluralism. Moreover, Evidential Pluralism favours the epistemic theory of causality over rival theories. Neither the difference-making nor mechanistic theories of causality seem to make sense of this kind of causal enquire in developmental psychology, because under a difference-making theory the search for mechanisms is inexplicable, and under a mechanistic theory the original focus on gaining evidence of correlation is inexplicable.

5.2 Neutrality in Cognitive Sciences

In line with the argument from neutrality presented in §2.4, we see next that by endorsing the epistemic theory of causality we can remain neutral about long-standing and intractable tensions in cognitive science. More precisely, we argue that by endorsing the epistemic theory we can, in some cases at least, remain neutral about how to individuate causally efficacious mental states.

To understand how endorsing the epistemic theory of causality can allow us to remain neutral about how to individuate causally efficacious mental states, it is helpful to consider an example: Marr's (1982) computational theory of vision. According to Marr's theory, perceptual mechanisms solve information-processing tasks set to them by nature with the aim of deriving a representation of three dimensional shape from information contained in two-dimensional images. Egan (1992, 453) gives a helpful recapitulation of this mechanism as follows:

> Marr's theory divides this task into three distinct stages, each involving the construction of a representation, tokens of which serve as inputs to subsequent processes. Vision culminates in a representation that is suitable for the recognition of objects. Innate assumptions [...] incorporated into the visual system itself, and reflecting physical constraints on the pairing of retinal images with distal shapes, allow the postulated mechanisms underlying early vision to recover information about the distal scene based only on information contained in the image.

To explain how the visual system undertakes this task, Marr's theory specifies a range of functions that are computed by the visual system—for example, the following function that characterises how the visual system initially filters the image:

$$\nabla^2 G * I(x, y) \tag{1}$$

Following Marr and Hildreth (1980), $\nabla^2 G$ is taken to represent a filter that "detect[s] intensity changes efficiently" in virtue, for example, of being:

> capable of being tuned to act at any desired scale, so that large filters can be used to detect blurry shadow edges, and small ones to detect sharply focused fine detail in the image (Marr, 1982, 54).[10]

$I(x, y)$ represents the image to be filtered, and $*$ the operation of convolution (which serves to determine the most important portions of an image) (Marr, 1982, 54-58, 338). The details here are complex, but all that matters is that, for Marr, this formal characterisation is, from a "computational point of view," a "precise specification of what the retina does" when initially filtering an image (see Figure 6).

Now, most cognitive theorists accept that Marr's computational explanation of vision is a causal explanation, but there has been heated debate about the locus of causality in this instance. Some argue that causal relations obtain between the contents of the representations over which the system computes; for example, that the system produces an early representation R_1 and the content of R_1 is causally efficacious in the production of later representations $R_2, \ldots, R_n$. Others, however, argue that causal relations obtain between physical properties of the system that realizes the functional (read: representational) states. Egan (1992, 446) takes this view when she argues that representations (as symbols):

> are just functionally characterized objects whose individuation conditions are specified by a realization function f_R which maps equivalence classes of physical features of a system to what we might call "symbolic" features. Formal operations are just those physical operations that are differentially sensitive to the aspects of symbolic expressions that under the realization function f_R are specified as symbolic features. The mapping f_R allows a causal sequence of physical state transitions to be interpreted as a *computation*.

At stake here is the question of whether or not the various modules of the visual system (i.e., what the system does) are "individuated essentially by reference to the contents of the representational tokens that form the inputs and outputs of these modules" (Egan, 1992, 453). Egan takes the view that they are not, because the realization function f_R should be understood as individuating computational states *non-semantically*. The standard view, however, is that such modules are individuated by reference to the contents of representational tokens, because intentional (read: semantic) mental states are individuated by their contents. Thus, there is tension between semantic and non-semantic individuations of the causally efficacious mental states involved in vision.

[10] According to Marr (1982, 55), $\nabla^2 G$ is a "circularly symmetric Mexican hat-shaped operator whose distribution in two dimensions may be expressed in terms of the radial distancer from the origin by the formula:

$\nabla^2 G(r) = -\frac{1}{\pi\sigma^4}\left(1 - \frac{r^2}{2\sigma^2}\right)\exp\left(\frac{r^2}{\sigma^2}\right)$"

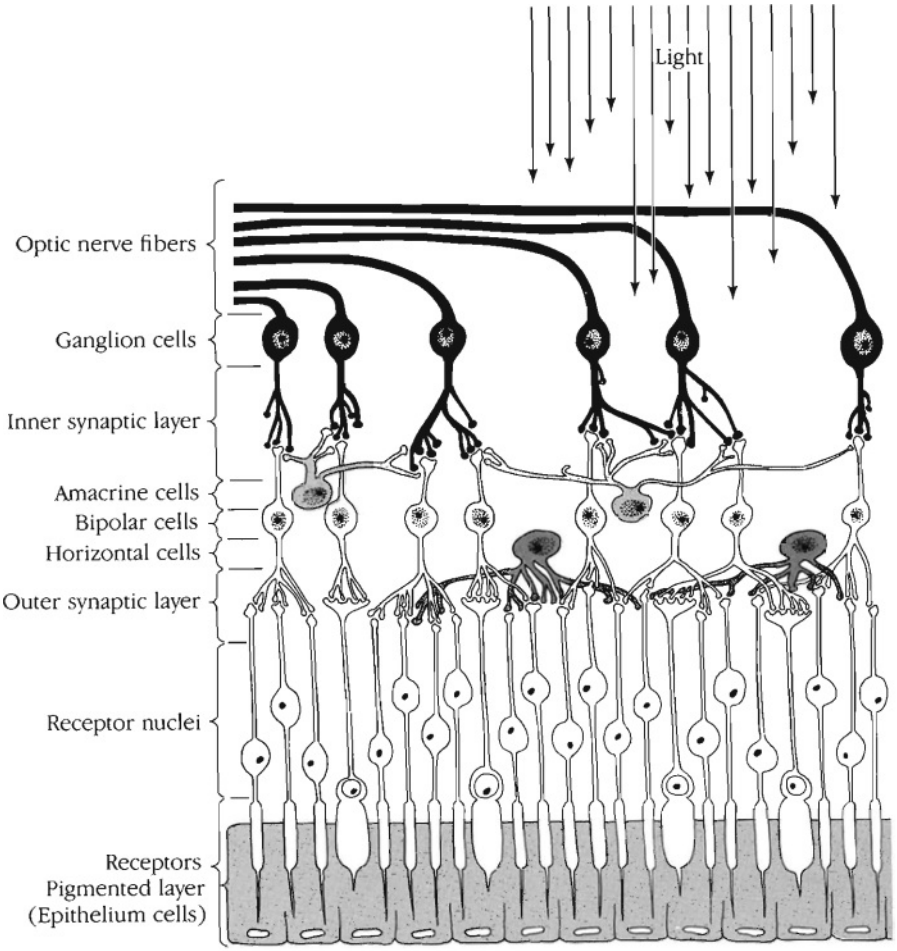

Figure 6: A cross section of the retina, part of whose function is to compute (1) from Marr, David. foreword by Shimon Ullman. afterword by Tomaso Poggio., *Vision*, Fig. 7.1 (p. 338), © 2010 Lucia M. Vaina, by permission of The MIT Press.

But this is not the only tension concerning the individuation of the causally efficacious mental states involved in vision, because even those who defend a semantic individuation of such states do not agree about whether the contents that individuate these states entirely supervene on intrinsic (physical) states of the subject possessing them. Some argue that the individuating contents do entirely supervene on intrinsic (physical) states of the subject possessing them (so-called internalists), but others argue that they do not (so-called externalists). This second tension, therefore, is about whether causally efficacious mental states are individuated (in part) by reference to the social and physical environment of the subject possessing them.

What is important here is that all can agree that Marr's computational theory of vision provides a good causal explanation of the cognitive competency in question (e.g., vision), while disagreeing about which mental states are causally

efficacious (non-semantic physical states vs representational states) or about the supervenience base of causally efficacious, representational mental states (intrinsic (physical) states vs (partly) social and physical environment). As a consequence, there is uncertainty about what exactly Marr's theory is a good causal explanation *of*: non-semantic physical states, representational states with contents supervening only on the intrinsic (physical) states of the subjects possessing them, or representational states with contents supervening (in part) on the social or physical environment.

Importantly, however, Marr argues that a good explanation of vision will incorporate descriptions at three "logically and causally related" levels: the level of "computational theory" (specifying the computed function), the level of representation and algorithm (describing how the function is computed), and the level of hardware implementation (describing the neural states supporting the computation). As such, Marr's theory is neutral about how to individuate the causally efficacious mental states of vision, because he asserts that a complete causal-explanation of vision will refer to both "neural mechanisms" at the level of hardware implementation and, at the level of computational theory, to "channels" in the "visual pathways" that detect spatial patterns "based on a form of spatial probability summation" (Marr, 1982, 62).[11]

And this is where the problem lies, because insofar as we endorse a difference-making or mechanistic theory of causality, we cannot remain neutral about how to individuate the causally efficacious mental states. To see why, note first that all of the possible interpretations of the relevant causally efficacious mental states can be said to stand in difference-making or mechanistic causal relations. For example, it is just as possible that there is an appropriate sort of mechanism linking non-semantic physical states as it is that there is an appropriate sort of mechanism linking representational states, and this holds true however you think the contents of representational states are individuated. Likewise, it is just as possible that, say, a chain of counterfactual dependence runs between non-semantic physical states as it is that a chain of counterfactual dependence runs between representational states.

The problem, however, is that mechanistic and difference-making theories of causality cannot remain neutral in this regard without incurring the charge of being vacuous. This is the case because it is only possible to spell-out the mechanistic or difference-making details when we take a view on how to individuate the causal relata, since the mechanisms or relations of difference-making will be different if we take these relata to be non-semantic physical states or representational states respectively. The key point, therefore, is that it is not possible for mechanistic and difference-making theories of causality to say simply that a computational state, A, caused some effect, B (e.g., the filtering of an image) without specifying how A is individuated, because then the relevant mechanism or difference-making relations are themselves left unspecified.

The epistemic theory of causality does not suffer from the same problems, because we need not take a view on how to individuate the relevant causally efficacious mental states in order to have a causal belief of the form $A \longrightarrow B$. In this way, the epistemic theory need not prejudge what is, ultimately, an open

[11] In fact, it is because Marr's theory is neutral in this way that philosophers have been able to disagree about how the causally efficacious mental states are to be individuated (cf. Burge, 1986; Egan, 1991; Kitcher, 1988; Segal, 1991).

question in the sciences of the mind. This follows because the epistemic theory of causality can cope with any possible individuation and, unlike mechanistic and difference-making theories of causality, need not take a stance on this issue while we remain unsure about which individuation is correct.[12] So, on the epistemic view, Marr's theory can be said to support a set of causal beliefs about causally efficacious computational states that allow us to predict, explain and control a particular portion of reality; namely, the cognitive system(s) responsible for vision. As such, we need only argue that Marr's theory supports causal beliefs of the form $A \longrightarrow B$, where A is, say, the transitions of those computational states (however they are individuated) carrying out function (1) above and B is the visual system's filtering of the image.

It is clear in practice that Marr's theory does support a range of causal beliefs of the form $A \longrightarrow B$ that motivate successful prediction, explanation, and control inferences. For example, the causal belief that our perceptual mechanisms (A_1) cause images processed on two separate retina to be "fused" if they occupy a region in visual space known as "Panum's fusional area" (B_1), which motivates the prediction that if images processed do not occupy Panum's fusional area, then we will lose the perception of these objects as being a single unified object and will see instead two images of the same object. Furthermore, the causal belief that disjunctive eye movement (A_2) causes changes in the plane of fixation (B_2), which motivates the control inference that if we prevent the eyes from moving disjunctively, then we prevent the visual system from changing the plane of fixation (by preventing the two lines of sight from converging or diverging).

It is clear that many working cognitive scientists do take Marr's theory to offer a causal explanation of vision, but this is obfuscated by the philosophical debate about how to individuate causally efficacious mental states. If we endorse the mechanistic or difference-making theories of causality, we cannot leave this debate behind. However, if we endorse the epistemic theory of causality, then we can agree that Marr's computational theory of vision provides a good causal explanation without getting bogged down in such debates. Thus, the epistemic theory will be appealing to those who are interested in defending a theory of causality that is able to prioritise scientific practice and consensus by remaining neutral about philosophical issues such as individuation.

6 Conclusion

Epistemic causality is a theory of the nature of causality, but one that gives primacy to the epistemology of causality. Causality is analysed in terms of rational causal beliefs, and the question of which causal beliefs are rational is one of causal epistemology. This tight connection between metaphysics, conceptual analysis and epistemology makes the epistemic theory an account of causality that is particularly close to scientific practice. We have seen that a careful consideration of both the social and cognitive sciences lends credibility to the epistemic causality. Our examples can be thought of as instantiating four arguments for epistemic causality: arguments from failure, parsimony, Evidential

[12] Of course, it might still matter in the end how the relevant causally efficacious mental states are individuated, because the causal belief $A \longrightarrow B$ might lead to successful PECs under one individuation but not under another.

Pluralism and neutrality. We appealed to the social sciences to illustrate the arguments from failure, parsimony and Evidential Pluralism and we appealed to the cognitive sciences to illustrate the arguments from Evidential Pluralism and neutrality.

Taking these arguments together with those of Russo and Williamson (2007) for epistemic causality in the health sciences, one can make the case that epistemic causality coheres well with causal practices across a broad range of sciences. This in itself is an important virtue of a theory of causality.

Acknowledgements

Some material from §4 appeared previously in Shan and Williamson (2023).

This research was supported by funding from the Leverhulme Trust (grants RPG-2019-059 and RPG-2022-336) and the British Academy (grant SRG1920/101076).

Bibliography

Adler, N. E., Boyce, T., Chesney, M. A., Cohen, S., Folkman, S., Kahn, R. L., and Syme, S. L. (1994). Socioeconomic status and health. The challenge of the gradient. *The American psychologist*, 49(1):15–24.

Adler, N. E. and Newman, K. (2002). Socioeconomic Disparities in Health: Pathways And Policies. *Health Affairs*, 21:60–76.

Ahn, W., Kalish, C. W., Medin, D. L., and Gelman, S. A. (1995). The role of covariation versus mechanism information in causal attribution. *Cognition*, 54(3):299–352.

Baddeley, A. (1992). Working memory: The interface between memory and cognition. *Journal of cognitive neuroscience*, 4(3):281–288.

Beach, D. and Pedersen, R. B. (2013). *Process-Tracing Methods*. The University of Michigan Press, Ann Arbor, MI.

Beebee, H. (2015). Causation, projection, inference and agency. In Johnson, R. N. and Smith, M., editors, *Passions and Projections: Themes from the Philosophy of Simon Blackburn*, pages 25–48. Oxford University Press.

Bosworth, B. (2018). Increasing Disparities in Mortality by Socioeconomic Status. *Annual Review of Public Health*, 39(1):237–251.

Burge, T. (1986). Individualism and psychology. *The Philosophical Review*, 95(1):3–45.

Carruthers, P. and Smith, P. K. (1996). *Theories of theories of mind*. Cambridge University Press.

Chater, N. and Vitányi, P. (2003). Simplicity: A unifying principle in cognitive science? *Trends in cognitive sciences*, 7(1):19–22.

Cohen, L., Lehéricy, S., Chochon, F., Lemer, C., Rivaud, S., and Dehaene, S. (2002). Language-specific tuning of visual cortex? functional properties of the visual word form area. *Brain*, 125(5):1054–1069.

Cook, T. D. and Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Rand McNally, Chicago, IL.

Crasnow, S. (2011). Evidence for Use: Causal Pluralism and the Role of Case Studies in Political Science Research. *Philosophy of the Social Sciences*, 41(1):26–49.

Cummins, R. C. (2000). "how does it work" versus "what are the laws?": Two conceptions of psychological explanation. In Keil, F. and Wilson, R. A., editors, *Explanation and Cognition*, pages 117–145. MIT Press.

Davidoff, J. (2001). Language and perceptual categorisation. *Trends in cognitive sciences*, 5(9):382–387.

Dehaene, S. (2009). *Reading in the brain: The new science of how we read*. Penguin.

Dehaene, S. and Cohen, L. (2007). Cultural recycling of cortical maps. *Neuron*, 56(2):384–398.

Dehaene, S., Le Clec'H, G., Poline, J., Le Bihan, D., and Cohen, L. (2002). The visual word form area: a prelexical representation of visual words in the fusiform gyrus. *Neuroreport*, 13(3):321–325.

Donohue, J. J. and Levitt, S. D. (2001). The impact of legalized abortion on crime. *The Quarterly Journal of Economics*, 116(2):379–420.

Dutton, D. B. (1986). Social class, health, and illness. In Aiken, L. and Mechanic, D., editors, *Applications of Social Science to Clinical Medicine and Health Policy*, pages 31–62. Rutgers University Press, New Brunswick, NJ.

Egan, F. (1991). Must psychology be individualistic? *The Philosophical Review*, 100(2):179–203.

Egan, F. (1992). Individualism, computation, and perceptual content. *Mind*, 101(403):443–459.

Evans, G. W. and Saegert, S. (2000). Residential Crowding in the Context of Inner City Poverty. In Wapner, S., Demick, J., Yamamoto, T., and Minami, H., editors, *Theoretical Perspectives in Environment-Behavior Research: Underlying Assumptions, Research Problems, and Methodologies*, pages 247–267. Springer US, Boston, MA.

Frith, U. and Happé, F. (1999). Theory of mind and self-consciousness: What is it like to be autistic? *Mind & language*, 14(1):82–89.

Gaillard, R., Naccache, L., Pinel, P., Clémenceau, S., Volle, E., Hasboun, D., Dupont, S., Baulac, M., Dehaene, S., Adam, C., et al. (2006). Direct intracranial, fmri, and lesion evidence for the causal role of left inferotemporal cortex in reading. *Neuron*, 50(2):191–204.

Gallagher, H. L. and Frith, C. D. (2003). Functional imaging of 'theory of mind'. *Trends in cognitive sciences*, 7(2):77–83.

Gerring, J. (2005). Causation: A unified framework for the social sciences. *Journal of Theoretical Politics*, 17(2):163–198.

Gerstenberg, T., Goodman, N. D., Lagnado, D. A., and Tenenbaum, J. B. (2022). A counterfactual simulation model of causal judgments for physical events. *Psychological Review*, 128(5):936–975.

Glynn, A. N. and Ichino, N. (2015). Using qualitative information to improve causal inference. *American Journal of Political Science*, 59(4):1055–1071.

Goertz, G. and Mahoney, J. (2012). *A tale of two cultures*. Princeton University Press, Princeton and Oxford.

Goldman, A. I. (1993). The psychology of folk psychology. *Behavioral and Brain sciences*, 16(1):15–28.

Goldthorpe, J. H. (2001). Causation, Statistics, and Sociology. *European Sociological Review*, 17(1):1–20.

Gopnik, A. (1993). How we know our minds: The illusion of first-person knowledge of intentionality. *Behavioral and Brain sciences*, 16(1):1–14.

Gopnik, A., Glymour, C., Sobel, D. M., Schulz, L. E., Kushnir, T., and Danks, D. (2004). A theory of causal learning in children: causal maps and Bayes nets. *Psychological Review*, 111(1):3–32.

Gould, E. D., Weinberg, B. A., and Mustard, D. B. (2002). Crime Rates and Local Labor Market Opportunities in the United States: 1979-1997. *The Review of Economics and Statistics*, 84(1):45–61.

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 37(3):424–438.

Granger, C. W. J. (1980). Testing for causality. *Journal of Economic Dynamics*, 2:329–352.

Guyatt, G., Cairns, J., Churchill, D., Cook, D., Haynes, B., Hirsh, J., Irvine, J., Levine, M., Levine, M., Nishikawa, J., Sackett, D., Brill-Edwards, P., Gerstein, H., Gibson, J., Jaeschke, R., Kerigan, A., Neville, A., Panju, A., Detsky, A., Enkin, M., Frid, P., Gerrity, M., Laupacis, A., Lawrence, V., Menard, J., Moyer, V., Mulrow, C., Links, P., Oxman, A., Sinclair, J., and Tugwell, P. (1992). Evidence-based medicine: A new approach to teaching the practice of medicine. *JAMA*, 268(17):2420–2425.

Haggard, S. and Kaufman, R. R. (2016). *Dictators and Democrats*. Princeton University Press, Princeton and Oxford.

Hall, N. (2004). Two concepts of causation. In Collins, J., Hall, N., and Paul, L., editors, *Causation and counterfactuals*, pages 225–276. MIT Press, Cambridge MA and London.

Harnad, S. (2017). To cognize is to categorize: cognition is categorization. In *Handbook of categorization in cognitive science*, pages 21–54. Elsevier.

Hedström, P. and Ylikoski, P. (2010). Causal Mechanisms in the Social Sciences. *Annual Review of Sociology*, 36(1):49–67.

Heinrich, J., Heine, S. J., and Norenzayan, A. (2010). The weirdest people in the world. *Behavioral and Brain Sciences*, 33(2-3):61–83.

Holland, P. W. (1986). Statistics and causal inference. *Journal of the American Statistical Association*, 81(396):945–960.

House, J. S., Lepkowski, J. M., Kinney, A. M., Mero, R. P., Kessler, R. C., and Herzog, A. R. (1994). The Social Stratification of Aging and Health. *Journal of Health and Social Behavior*, 35(3):213–234.

Hume, D. (1748). *An enquiry concerning human understanding*. Oxford Philosophical Texts. Oxford University Press, Oxford.

Illsley, R. and Mullen, K. (1985). The health needs of disadvantaged client groups. In Holland, W. W., Detels, R., and Knox, G., editors, *Oxford Textbook of Public Health*, pages 389–402. Oxford University Press, Oxford.

Imai, K. (2017). *Quantitative Social Science: An Introduction*. Princeton University Press, Princeton.

Jaynes, E. T. (2003). *Probability theory: the logic of science*. Cambridge University Press, Cambridge.

Kaplan, D. and Craver, C. F. (2011). The explanatory force of dynamical and mathematical models in neuroscience: A mechanistic perspective. *Philosophy of science*, 78(4):601–627.

King, G., Keohane, R. O., and Verba, S. (1994). *Designing social inquiry*. Princeton University Press, Princeton, NJ.

Kitcher, P. (1988). Marr's computational theory of vision. *Philosophy of Science*, 55(1):1–24.

Lagnado, D. A. and Sloman, S. (2004). The advantage of timely intervention. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 30(4). To appear.

Lantz, P. M., House, J. S., Lepkowski, J. M., Williams, D. R., Mero, R. P., and Chen, J. (1998). Socioeconomic factors, health behaviors, and mortality: results from a nationally representative prospective study of US adults. *Journal of the American Medical Association*, 279(21):1703–1708.

Leslie, A. M. (1987). Pretense and representation: The origins of" theory of mind.". *Psychological review*, 94(4):412.

Leslie, A. M., Friedman, O., and German, T. P. (2004). Core mechanisms in 'theory of mind'. *Trends in cognitive sciences*, 8(12):528–533.

Levitt, S. D. (2004). Understanding Why Crime Fell in the 1990s: Four Factors that Explain the Decline and Six that Do Not. *Journal of Economic Perspectives*, 18(1):163–190.

Lewis, D. (1973). Causation. *Journal of Philosophy*, 70:556–567.

Lin, E. L. and Murphy, G. L. (1997). Effects of background knowledge on object categorization and part detection. *Journal of Experimental Psychology: Human Perception and Performance*, 23(4):1153–1169.

Link, B. G. and Phelan, J. C. (1995). Social Conditions As Fundamental Causes of Disease. *Journal of Health and Social Behavior*, (Extra issue):80–94.

Longworth, F. (2006). Causation, pluralism and responsibility. *Philosophica*, 77:45–68.

Lutfey, K. and Freese, J. (2005). Toward Some Fundamentals of Fundamental Causality: Socioeconomic Status and Health in the Routine Clinic Visit for Diabetes. *American Journal of Sociology*, 110(5):1326–1372.

Machin, S. and Meghir, C. (2004). Crime and Economic Incentives. *The Journal of Human Resources*, 39(4):958–979.

Marr, D. (1982). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information.* New York: W. H. Freeman.

Marr, D. and Hildreth, E. (1980). Theory of edge detection. *Proc. R. Soc. Lond B*, 207:187–217.

Maziarz, M. (2020). *The philosophy of causality in economics: Causal inferences and policy proposals.* Routledge, London and New York.

Maziarz, M. (2021). Resolving empirical controversies withe mechanistic evidence. *Synthese*, 199:9957–9978.

Meyer, R. (2020). The non-mechanistic option: Defending dynamical explanations. *The British Journal for the Philosophy of Science*, 71(3):959–985.

Michaelian, K. and Sutton, J. (2013). Distributed cognition and memory research: History and current directions. *Review of philosophy and psychology*, 4(1):1–24.

Morgan, S. L. and Winship, C. (2015). *Counterfactuals and Causal Inferences: Methods and Principles for Social Research.* Cambridge University Press, Cambridge, 2 edition.

Morrison, A. B. and Chein, J. M. (2011). Does working memory training work? the promise and challenges of enhancing cognition by training working memory. *Psychonomic bulletin & review*, 18(1):46–60.

Onishi, K. H. and Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science*, 308(5719):255–258.

Pampel, F. C., Krueger, P. M., and Denney, J. T. (2010). Socioeconomic Disparities in Health Behaviors. *Annual review of sociology*, 36:349–370.

Parkkinen, V.-P., Wallmann, C., Wilde, M., Clarke, B., Illari, P., Kelly, M. P., Norell, C., Russo, F., Shaw, B., and Williamson, J. (2018). *Evaluating evidence of mechanisms in medicine: principles and procedures*. Springer, Cham, Switzerland.

Patton, M. Q. (1990). *Qualitative Research and Evaluation Methods*. Sage, Thousand Oaks, CA, third (2002) edition.

Phelan, J. C., Link, B. G., Diez-Roux, A., Kawachi, I., and Levin, B. (2004). "Fundamental causes" of social inequalities in mortality: A test of the theory. *Journal of Health and Social Behavior*, 45:265–285.

Phelan, J. C., Link, B. G., and Tehranifar, P. (2010). Social conditions as fundamental causes of health inequalities: theory, evidence, and policy implications. *Journal of health and social behavior*, 51 Suppl:S28–S40.

Piccinini, G. and Craver, C. (2011). Integrating psychology and neuroscience: Functional analyses as mechanism sketches. *Synthese*, 183(3):283–311.

Premack, D. and Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and brain sciences*, 1(4):515–526.

Reich, L., Szwed, M., Cohen, L., and Amedi, A. (2011). A ventral visual stream reading center independent of visual experience. *Current Biology*, 21(5):363–368.

Reiss, J. (2009). Causation in the social sciences: Evidence, inference, and purposes. *Philosophy of the Social Sciences*, 39(1):20–40.

Reiss, J. (2012). Causation in the sciences: An inferentialist account. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 43(4):769–777. Causality in the Biomedical and Social Sciences.

Rohlfing, I. and Zuber, C. I. (2021). Check Your Truth Conditions! Clarifying the Relationship between Theories of Causation and Social Science Methods for Causal Inference. *Sociological Methods and Research*, 50(4):1623–1659.

Rothman, K. (1986). *Modern Epidemiology*. Little, Brown, and Company, Boston, MA.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66:688–701.

Russo, F. (2009). *Causality and causal modelling in the social sciences: measuring variations*. Springer.

Russo, F. and Williamson, J. (2007). Interpreting causality in the health sciences. *International Studies in the Philosophy of Science*, 21(2):157–170.

Russo, F. and Williamson, J. (2011a). Epistemic causality and evidence-based medicine. *History and Philosophy of the Life Sciences*, 33(4):563–582.

Russo, F. and Williamson, J. (2011b). Generic versus single-case causality: the case of autopsy. *European Journal for Philosophy of Science*, 1(1):47–69.

Sackett, D. L., Rosenberg, W. M. C., Gray, J. A. M., Haynes, R. B., and Richardson, W. S. (1996). Evidence based medicine: what it is and what it isn't. *British Medical Journal*, 312(7023):71–72.

Schnall, P. L., Pieper, C., Schwartz, J. E., Karasek, R. A., Schlussel, Y., Devereux, R. B., Ganau, A., Alderman, M., Warren, K., and Pickering, T. G. (1990). The relationship between 'job strain,' workplace diastolic blood pressure, and left ventricular mass index. Results of a case-control study. *Journal of the American Medical Association*, 263(14):1929–1935.

Schurz, M. and Perner, J. (2015). An evaluation of neurocognitive models of theory of mind. *Frontiers in psychology*, 6:1610.

Segal, G. (1991). Defence of a reasonable individualism. *Mind*, 100(4):485–494.

Shan, Y. and Williamson, J. (2021). Applying Evidential Pluralism to the social sciences. *European Journal for the Philosophy of Science*, 11(4):96.

Shan, Y. and Williamson, J. (2023). *Evidential Pluralism in the social sciences*. Routledge, London and New York.

Sims, C. R. (2018). Efficient coding explains the universal law of generalization in human perception. *Science*, 360(6389):652–656.

Sobel, D. M. and Kushnir, T. (2006). The importance of decision making in causal learning from interventions. *Memory & Cognition*, 34(2):411–419.

Southgate, V., Senju, A., and Csibra, G. (2007). Action anticipation through attribution of false belief by 2-year-olds. *Psychological science*, 18(7):587–592.

Surian, L., Caldi, S., and Sperber, D. (2007). Attribution of beliefs by 13-month-old infants. *Psychological science*, 18(7):580–586.

Tanrıkulu, Ö. D., Chetverikov, A., Hansmann-Roth, S., and Kristjánsson, Á. (2021). What kind of empirical evidence is needed for probabilistic mental representations? an example from visual perception. *Cognition*, 217:104903.

Taylor, S. D. (2021). Causation and cognition: An epistemic approach. *Synthese*, 199:9133–9160.

Taylor, S. D. and Sutton, P. R. (2021). A frame-theoretic model of bayesian category learning. In *Concepts, frames and cascades in semantics, cognition and ontology*, pages 329–349. Springer, Cham.

Teddlie, C. and Tashakkori, A. (2009). *Foundations of Mixed Methods Research: Integrating Quantitative and Qualitative Approaches in the Social and Behavioral Sciences*. Sage, Thousand Oaks, CA.

Thagard, P. (2005). *Mind: Introduction to cognitive science*. Cambridge, MA: MIT press.

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, Cambridge.

Wilde, M. and Williamson, J. (2016). Evidence and epistemic causality. In Wiedermann, W. and von Eye, A., editors, *Statistics and Causality: Methods for Applied Empirical Research*, pages 31–41. Wiley, Hoboken, New Jersey.

Williamson, J. (2005). *Bayesian nets and causality: philosophical and computational foundations*. Oxford University Press, Oxford.

Williamson, J. (2006a). Causal pluralism versus epistemic causality. *Philosophica*, 77:69–96.

Williamson, J. (2006b). Dispositional versus epistemic causality. *Minds and Machines*, 16:259–276.

Williamson, J. (2007). Causality. In Gabbay, D. and Guenthner, F., editors, *Handbook of Philosophical Logic*, volume 14, pages 95–126. Springer, Dordrecht.

Williamson, J. (2009). Probabilistic theories. In Beebee, H., Hitchcock, C., and Menzies, P., editors, *The Oxford Handbook of Causation*, pages 185–212. Oxford University Press, Oxford.

Williamson, J. (2011). Mechanistic theories of causality. *Philosophy Compass*, 6(6):421–447.

Williamson, J. (2013). How can causal explanations explain? *Erkenntnis*, 78:257–275.

Williamson, J. (2021). Calibration for epistemic causality. *Erkenntnis*, 86(4):941–960.

Williamson, J. (2022). A Bayesian account of establishing. *The British Journal for the Philosophy of Science*, 73(4):903–925.

Woodward, J. (2003). *Making Things Happen: A Theory of Causal Explanation*. Oxford University Press, Oxford.