

On pain experience, multidisciplinary integration and the level-laden conception of science

Tudor M. Baetu¹

Received: 30 October 2016 / Accepted: 4 May 2017 / Published online: 23 May 2017
© The Author(s) 2017. This article is an open access publication

Abstract Multidisciplinary models aggregating ‘lower-level’ biological and ‘higher-level’ psychological and social determinants of a phenomenon raise a puzzle. How is the interaction between the physical, the psychological and the social conceptualized and explained? Using biopsychosocial models of pain as an illustration, I argue that these models are in fact level-neutral compilations of empirical findings about correlated and causally relevant factors, and as such they neither assume, nor entail a conceptual or ontological stratification into levels of description, explanation or reality. If inter-level causation is deemed problematic or if debates about the superiority of a particular level of description or explanation arise, these issues are fueled by considerations other than empirical findings.

Keywords Philosophy of biology · Experimental research · Pain · Causality

1 Multidisciplinary models

The diagnosis and pathology of many psychiatric conditions encompasses a mixture of physiological and psychological symptoms, risk factors and causal determinants (Kendall 2001; Kendler and Campbell 2009; Murphy 2006; Schaffner 1993). Epidemiological studies reveal that many common diseases are multifactorial and can be influenced by factors such as genes, diet, life style and socioeconomic status (Broadbent 2013, Ch. 10; Howick 2011). Biopsychosocial models of somatic disease emphasize multiple levels of medical intervention (Engel 1977, 1980).

✉ Tudor M. Baetu
tudor.baetu@bristol.ac.uk

¹ Department of Philosophy, Cotham House, University of Bristol, Bristol BS6 6JL, UK

Finally, in basic science, there are many examples of models of psychological and social phenomena which include biological determinants and sometimes cases where psychological and social factors make an appearance in biological explanations.

To give a particularly striking example, one of the most revolutionary features of the gate control model of pain lies in the attempt to integrate neurophysiological and psychological determinants of pain experience. The model posits that a neural circuit mechanism in the dorsal horn of the spinal cord modulates signals from nociceptors by integrating inputs from thermo-mechanical receptors and from brain areas associated with cognitive and emotional appraisals, thus altering the quality and intensity of pain experience (Melzack and Wall 1965). The proposal opened the door to multidisciplinary approaches yielding a variety of biopsychosocial models of pain according to which pain experience “is determined by the interaction among biological, psychological (which include cognition, affect, behavior), and social factors (which include the social and cultural contexts that influence a person’s perception of and response to physical signs and symptoms)” (Asmundson and Wright 2004, 42).

Yet, despite their popularity, multidisciplinary models raise a puzzle. Each discipline provides its own unique perspective on a given phenomenon which is insulated from other disciplinary perspectives. Some disciplinary boundaries are historical contingencies or by-products of social and political agendas. In such cases, multidisciplinary research is made possible by a restructuring of scientific institutions. Other boundaries, however, are taken to reflect more profound epistemic and metaphysical differences, such as differences in experimental methods, distinct theoretical frameworks and even divergent assumptions about the nature of reality. These differences, which cannot be removed by a mere reshuffling of institutional affiliations and funding incentives, motivate much stronger ‘level’ distinctions which seem to undermine multidisciplinary research.

In this paper, I argue that multidisciplinary models reveal an important, but underestimated feature of empirical research, namely its neutrality in respect to levels of description, explanation and reality. The upshot of this claim is that if debates about the epistemic superiority of a particular level of description or explanation arise, or if ontological distinctions are drawn between intra- and inter-level causation, somatic and psychogenic etiology or pathogenesis, then these issues are due to considerations other than empirical research.

In Sect. 2, I discuss the rationales for distinguishing levels of description, explanation and reality, and explain how a level-laden conception of science clashes with multidisciplinary research. I argue that the clash is resolved either by assuming the eventual reduction of psychological and social perspectives to biological explanations; or, more interestingly, by engaging in a largely experimental research program aiming to generate and systematize findings about empirically established correlations and causal determinants of a particular phenomenon.

In Sect. 3, I characterize biopsychosocial models of pain in particular and multidisciplinary models in general as attempts to summarize what is known about correlates and causal-mechanistic determinants of a given phenomenon. Such models can be viewed as partial explanations pointing to causes and mechanisms of phenomena, can

be primarily descriptions of phenomena or of correlations associated with phenomena, or may involve a mixture of both.¹ I argue that these models can aggregate any correlate or causal determinant of a phenomenon of interest, be it biological, psychological, social or other, in virtue of a methodological principle of epistemic parity among the investigative methodologies driving experimental and clinical research across disciplines. This principle denies empirical justification of the exclusion of some factors from the model, the segregation of certain kinds of factors into distinct models, or the layering of factors included in any given model along two or more levels.²

In Sect. 4, I address two objections to the level-neutrality thesis. First, I address the possibility that different levels of epistemic confidence are systematically assigned to findings from different disciplines based on differences in the validity, reliability and accuracy of the methods they employ. Such differences would justify an experimental reductionism whereby the methods of some disciplines are systematically replaced by methods from other disciplines, ultimately leading to a preferred level of explanation. My response is that the methods of ‘lower-level’ disciplines (e.g., biological methods of assessment of subjective experience) cannot systematically replace the methods of ‘higher-level’ disciplines (e.g., psychological methods) because the latter are needed to establish the validity of the former. The second objection is a version of Kim’s argument from causal exclusion. It seems reasonable to assume that any given multidisciplinary model will ultimately grow to encompass several layers of causes each seemingly sufficient to account for the same range of effects on the phenomenon to be explained. I address the objection by drawing a distinction between evidence for causal relevance and evidence for the completeness, or causal sufficiency of an explanation. I argue that while causal relevance is enough to warrant supervenience, it does not support conclusive inferences about causal sufficiency. Without causal sufficiency, there is no reason to assume causal exclusion.

In Sect. 5, I conclude that we must seek the rationales for accounts and debates involving levels of explanation and reality elsewhere than empirical research. I propose that, while empirical research has the internal resources for sustaining integrative multidisciplinary research, level distinctions reflect the fact that different disciplines rely on local theoretical frameworks which fail to add up to a coherent unified theory capable of accounting for all the discoveries of empirical research.

¹ The term ‘model’ has many understandings in life sciences alone (Baetu 2014; Bolker 2009; Leonelli 2007), in addition to numerous other characterizations found in other scientific disciplines, philosophy of science, logic and mathematics. I focus exclusively on multidisciplinary models as characterized above.

² The notion that empirical claims are not laden by metaphysical assumptions about levels of reality is by no means new. Logical positivists and, later on, proponents of the identity theory of the mind argued that the language scientists use to describe empirical findings is topic-neutral (Feigl 1967; Smart 1959). Unfortunately, this strategy collapsed into a debate about whether the distinctive terminologies of biology and psychology reflect significant assumptions about differences in the nature of the factors involved (Rosenthal 1994). What is proposed here is a new approach switching the focus from the language used to describe empirical claims to the methods used to obtain it.

2 The puzzle of multidisciplinary research

2.1 The level-laden conception of science

A distinction is commonly drawn between a physical and a phenomenological, or empirical, level of description, where the former is studied by ‘lower-level’ explanatory sciences such as molecular biology, while the latter belongs to psychology and other ‘higher-level’ descriptive disciplines, such as zoology and embryology. Descriptive sciences tend to rely on direct observations of macro-variables in order to generate classifications, generalizations and predictions. Explanatory sciences, on the other hand, introduce a host of micro-entities whose existence can only be indirectly corroborated by sophisticated experimental protocols. The boundaries separating levels of description can be quite vague, as it is not always clear what counts as direct and indirect observation, or whether empirical generalizations are devoid of any explanatory potential. Nevertheless, the fact that micro-entities are often times hypothesized in order to provide reductive explanations of the relationships between macro-variables tends to maintain a separation between the two, the latter constituting the domain of phenomena to be explained, the former the domain of whatever does the explaining.

While multidisciplinary research may find many natural niches along the poorly defined borders separating levels of description, there are cases where levels morph into much stronger distinctions. For example, somatic diseases and biological explanations are still distinguished from mental diseases and psychological explanations based on theoretical expectations of what can and cannot interact with the components of a physiological mechanism. This interaction criterion has been used to reject the notion that a proper explanation can mix neurological and psychological considerations (Frith 1992, Ch. 3), demand that one must have an appropriate account of how they are related to one another (Kim 1993), and criticize multifactorial models of disease aggregating biological and socioeconomic risk factors (Gori 1989; Krieger 1994). In all these cases, the argument is that psychological and social factors cannot simply be ‘plugged’ into biological mechanisms, but must first be shown to have a biological basis which may then be integrated in the biological machinery. As a result, a sharp distinction is drawn between ‘lower-level’ explanations involving biological mechanisms and the ‘higher-level’ findings, descriptions and explanations of psychology and sociology.³

Levels of explanation usually go hand in hand with levels of reality. The fact that, in many cases, it is possible to manipulate micro-entities in order to generate macro-level changes supports an argument for realism about micro-entities (Hacking 1983, Ch. 16). This interpretation is widely accepted in biology, which takes biological mechanisms, including cellular and molecular ones, to be real things rather than mere explanatory constructs (Craver and Darden 2013, Ch. 6). A realist interpretation of the levels of explanation conception gives us a ‘soft’ dualist ontology according to

³ Note that there is nothing inherent to mechanistic explanation that assumes or entails such a stratification. The stratification is the consequence of an independent requirement that the interactions between the parts of a mechanism must be of a particular type (e.g., mechanical, chemical). This requirement is not assumed in psychology, sociology and economics, where it is not uncommon to find references to social mechanisms consisting of individuals interacting according to certain rules.

which there is a phenomenological/empirical and a physical level of reality. While the latter is assumed to be dependent on the former, the interaction criterion introduced earlier prohibits the mixing together of elements belonging to the two levels in the same mechanism. This promotes a strictly regimented segregation: elements belonging to the ‘lower’ physical level are components of biological mechanisms, while the elements belonging to the ‘higher’ phenomenological level correspond to descriptions of the phenomena for which these mechanisms are responsible. This kind of dichotomy is often times assumed when mental diseases and biological explanations are distinguished from somatic diseases and psychological explanations (Kendler and Campbell 2009). In a more drastic version, it is argued that the phenomenological level of conscious experience and the biological level of mechanistic explanation are separated by an unbridgeable explanatory gap which forces us either to accept a dualist position contrasting external reality to an irreducible subject-bounded mental reality (Chalmers 1996, Ch. 4; Levine 1983), or to eliminate the phenomenological in favor of the physical (Dennett 1996, Ch. 11).

2.2 Integration and multidisciplinary research

If we accept this level-laden conception of science,⁴ it is not clear what kind of integrative work can be done by multidisciplinary models. Either these models specify how psychological and social factors interact with biological mechanisms, or they are nothing else but a disjoint patchwork of incompatible perspectives. Multidisciplinarity may be understood along the lines of a pragmatic pluralism alternating between distinct disciplinary perspectives, but this is not the same as a novel, genuinely integrative perspective.

If pressed on the issue, researchers often point in the direction of yet to be discovered “biological substrates of specific psychological processes” (Craig and Versloot 2011, 27). The suggestion here is that the explanatory gap between the biological and the psychological will eventually be closed, most likely in the reductive sense that psychological determinants will be someday be accountable in biological terms. However, if we resist for a moment the temptation of reductionism as a path out of the difficulty, we soon discover that there is a second answer available in the scientific literature. If we ask “What can possibly bring together considerations as remote as ‘increased nociceptor activity’, ‘catastrophizing’ and ‘spousal support’?”, the answer we are given is that all three are causal determinants of pain experience specifying “targets for different forms of therapeutic intervention” (Craig and Versloot 2011, 27). In other words, it is possible to alter pain experience by experimentally intervening on any of these determinants, irrespective of whether we may think of them as being ‘lower’ or ‘higher-level’. Furthermore, some meta-analysis studies suggest that multimodal

⁴ I explicitly exclude from this conception levels of composition, which can be understood in strictly empirical terms (Bechtel and Richardson 2010, Chs. 3–4; Wimsatt 1972). For instance, brains are experimentally decomposable into neurons and neural pathways, while pain experience can be experimentally dissociated into a sensory-discriminative component capturing the intensity, location, quality and duration of the pain, and an affective-motivational component referring to the unpleasantness of the experience. By themselves, levels of composition don’t entail any explanatory gaps or ontological discontinuities in the fabric of reality.

therapies simultaneously targeting several determinants, for instance by combining pharmaceutical interventions with cognitive-behavioral therapy and social environment interventions, can be more effective than single-level interventions (Rossy et al. 1999). This indicates that factors we may take to belong to distinct levels can interfere with each other's ability to alter pain experience.

What is particularly appealing about this second answer is that, unlike the yet to be achieved goals of reductionism, it is firmly grounded in the experimental evidence available today. From the standpoint of empirical research, multidisciplinary models do genuine integrative work, albeit not by providing a unified explanation of how psychosocial factors interact with biological mechanisms, but rather by acknowledging evidence that the phenomenon of pain is causally determined by both biological and psychosocial factors and by attempting to summarize what is known about these factors and their interactions. In the remainder of the paper, I develop and defend a methodological argument justifying the aggregation of empirical findings irrespective of the level distinctions that may separate the disciplines in which these findings have been generated.

3 A methodological argument for level-neutrality

3.1 The interventionist account of causation

According to an interventionist account of causation, “causal (as opposed to merely correlational) relationships are relationships that describe what will happen to some variables (effects) when we manipulate or intervene on others (causes). To say that a relationship is causal is to say that it is exploitable for purposes of manipulation and control in a way that merely correlational relationships are not” (Woodward 2008). Ideally, evidence for causation requires that an intervention on causal factor X must change the outcome Y without changing any other variable that is a cause of Y —that is, without directly changing Y , or any other variable along the causal pathway from X to Y , or by simultaneously intervening on convergent causal pathways leading to Y (Woodward 2003, pp. 94–99); this is meant to ensure that X , and not some other variable, is the difference maker targeted by the intervention.

The account captures the distinction between experiments and observational studies. For example, an epidemiological study may demonstrate that, in a given population, a particular DNA sequence is associated with a given phenotype in a statically significant way. We may conclude that the two are strongly correlated, but there is no conclusive evidence for causation. Two main problems arise with strictly observational studies. First, there is no control group to demonstrate that genotype and phenotype are not effects of a common causal background. Second, there is nothing to support the claim that the genotype causes the phenotype and not the other way around. Association is inconclusive evidence for causation in the sense that it does not allow us to discriminate between several interpretations, of which only one corresponds to the desired causal claim.

In contrast, an experimental intervention can provide conclusive evidence for causation because it allows us to discriminate and reject alternative interpretations of the

results. For example, a gene knockout experiment may reveal that if the DNA sequence is mutated or deleted, then the phenotype is altered; this establishes order of causation. Given that the change in phenotype does not happen in the genetically unmodified, but otherwise identical control system, we may conclude that the difference maker is the DNA sequence and not another cause. In the converse experiment, when the sequence is restored, the phenotype is restored. Again, the temporal sequence is clear, and given that the genetically modified system does not recover on its own, we can again conclude that the relevant difference maker is that sequence and not some other cause.

3.2 The interventionist argument for the level-neutrality of causal explanations

James Woodward argues that, in the context of an interventionist account of causation, nothing justifies giving “automatic or a priori preference to any particular grain or ‘level’ of causal description over any other”; furthermore, “there is no bar in principle to mixing variables that are at what might seem to be different ‘levels’ in causal claims” (2008, 222). The argument rests on a methodological consideration: if an interventionist account is all that is needed to justify claims about causation in empirical science, and if the same method of ideal interventions is successfully applied in all cases, then there is no reason why some causal determinants should be assigned a distinct status. Causal models are level-neutral in the sense that they can amalgamate any kind of determinant that satisfies the desiderata of a controlled intervention experiment.

This realization led [Woodward \(2008\)](#), [Campbell \(2008, 2013\)](#) and ([Kendler and Campbell 2009](#)) to further argue that in some fields of investigation, such as psychiatry, there are no preferred levels of explanation favoring particular kinds of determinants. This conclusion rests on the fact that an interventionist account of causation provides at the same time an account of experimental methodology, namely the desiderata a study must satisfy in order to demonstrate causation, as well as an account of scientific explanation, where explanations are construed as answers to ‘what-if-things-had-been-different’ questions ([Woodward 2003](#), Ch. 5).

3.3 Disentangling methodology and explanation

According to the interventionist argument, level-neutrality is a peculiarity of explanations in disciplines where experimental results that satisfy the desiderata of interventionist approaches are accepted as satisfactory explanations marking the end-point of scientific inquiry. This is rather restrictive. Causal models are explicitly recognized as end-goals of scientific inquiry only in the context of clinically oriented disciplines and approaches, such as epidemiology and evidence-based medicine, notorious for their pragmatic emphasis on ‘what works in practice’. They are not the gold standard of scientific explanation in science, as reflected by the hypotheses driving most biomedical research or the kind of explanations one typically finds in science textbooks. Yet, despite this qualification, it would be wrong to conclude that causal models are absent or less important outside clinical research. Quite on the contrary, causal models are widespread across sciences and play a crucial role in

discovery, experimentation and the development of medical and technological applications. Causal models are extremely common in basic science, albeit not as a form of explanation, but rather as attempts to summarize what is currently known about the causal determinants of a phenomenon.

It is therefore desirable to disentangle claims about experimental methodology from claims about what counts as an acceptable scientific explanation. Focusing on experimental methodology has two important benefits. First, there is a gain in generality. An interventionist account of causation captures a widespread commitment to experimental research going well beyond the explanatory ideals of clinical research. Thus, level-neutrality becomes a general feature of causal models systematizing experimental results, rather than remaining confined to a particular set of disciplines and their explanations. Second, there is a shift from explanation to empirical findings. While the level-neutrality of causal models as explanations is interesting because it reveals the non-reductive character of explanations typically associated with clinical sciences, the level-neutrality of causal models as compilations of empirical findings is interesting because it reveals an important way in which empirical research is independent from theoretical and metaphysical assumptions about levels of explanation and reality. An important consequence of this independence is the ability to integrate findings under the methodological umbrella of empirical research whether or not currently available theoretical frameworks can keep up with the demands of such integration.

3.4 Level-neutrality as a general feature of empirical research

With the explanatory commitments of interventionism out of the way, we can now inspect more closely the core methodological argument behind the level-neutrality thesis. The argument relies on two premises: (1) the different causes included in a model target the same effect; and (2) the different causes must be established with a similar degree of epistemic certainty. However, it turns out that these two premises can also be used to argue that experimental results in general can be combined in a level neutral way. Therefore, under the assumption that the interventionist argument for level-neutrality is sound, level-neutrality holds for experimental results in general.⁵ I will begin by discussing the first premise in this section, and continue with a discussion of the second premise in Sect. 3.5.

The first thing worth noting is that interventionist desiderata are insufficient to justify the aggregation of causal factors in the same model. One can easily imagine a scenario where two distinct experiments involving ideal interventions demonstrate the causal relevance of a biological and a psychological factor to pain intensity, but in each case pain is measured in a different way. In such a scenario, it is not immediately obvious that we can legitimately aggregate the two determinants in the same model. Interventionist criteria only assess evidence for causal relationships. If the same method of ideal interventions is applied to a psychological and a biological factor and a measurable effect is observed in each case, then we can conclude that the former is

⁵ I am indebted to an anonymous reviewer for this concise formulation of the argument.

as much as a cause as the latter. However, this tells us nothing about both being causes of the same effect.

Something else is required, namely evidence that different findings are about the same phenomenon. This is achieved by means of standardization (e.g., standardized questionnaires, genetically uniform organisms) and operationalization (the step-by-step instructions detailing experimental procedures in the ‘Materials and Methods’ section of scientific articles). These practices specify the exact nature of interventions and increase the likelihood that the test and control conditions are identical except for the variables under direct intervention, as required by interventionism. However, they also aim to ensure that phenomena, objects of study, experimental setups and methods of investigation can be identified and replicated by different research teams (Ankeny 2001; Gossel 1992; Müller-Wille 2007; Weber 2008). In turn, findings generated in the context of a standardized and operationalized experimental setup are more likely to be about ‘the same thing’, and therefore can be directly compared and aggregated in the same model (Baetu 2014).⁶

The implication here is that what glues together the various elements of a model is not evidence for causality simpliciter, but evidence for causal relevance with respect to the same phenomenon. This suggests that level-neutrality is not a unique feature of causal models, but rather a general feature of models broadly construed to include descriptions of phenomena, webs of correlated factors and causal determinants. For example, just as the causal contributions of inflammatory responses and anticipation to pain experience are combined in the same causal model of chronic pain, statistically significant neural and cultural correlates of pain experience are jointly considered as empirical constraints a successful explanation of pain should take into consideration, while physiological and psychological symptoms are aggregated in the same description of the phenomenon of pain.

The argument can be further generalized. Models may also mix correlated and causal determinants, along with occasional bits and pieces of better understood mechanistic details.⁷ The key requirement is that they are all correlates, causes and mechanistic details of the same phenomenon. This is consistent with a well-documented sequence of events in the lifecycle of many research projects in biomedical sciences, namely a gradual progression from an initial fixing of explananda as descriptions of phenomena to an expansion into wider webs of correlated factors (Bogen and Woodward 1988; Leonelli 2009; McAllister 1997), which in turn provide a pool of putative determinants subsequently tested for causal relevance (Baetu 2012; Craver

⁶ In biomedical sciences, empirical findings are often aggregated in the same model in virtue of extrapolations to systems and situations others than those actually assessed in a study. In such cases, it is important to establish the external validity of experimental results (Baetu 2016a; Germain and Baetu 2017; Steel 2007).

⁷ The latter amount to ‘higher resolution’ descriptions of causal determinants and their interactions. Mechanistic details may include information about the spatiotemporal organization of causal pathways (compartmentalization, differential gradients, temporal dynamics), the nature of causal interactions (geometrical fit, phosphorylation, diffusion, electron exchange), and functional role ascriptions based on how a phenomenon changes qualitatively and quantitatively in response to various types of interventions on causal determinants. These details typically rely on an uncontroversial interpretation of experimental results, and are included in the model strictly on the basis of their causal relevance.

2007, Chs. 2–4; Woodward 2002). Ultimately, knowledge of causal determinants provides the materials for hypothesizing mechanistic explanations, as well as a set of empirical constraints that limits the space of possible mechanisms and against which hypothesized mechanisms can be tested (Craver and Darden 2013, Chs. 7–9; Darden 2006, Ch. 12).

On this account, multidisciplinary models are expected to fall on a more or less continuous spectrum between descriptions of phenomena and mechanistic explanations. This may seem to include a lot of things, but in the end the difference between these elements hinges on experimental limitations. Any clinical study aims to demonstrate causation, but often times can only gather evidence for correlation. In basic science, the gold standard is detailed mechanistic explanation providing a step by step description of how an organized set of entities interact in order to produce a phenomenon of interest, but the experimental evidence may only allow for the reconstruction of rough causal pathways surrounded by a halo of potentially relevant correlates whose exact mechanistic roles are still unknown.

This description matches the composition of biopsychosocial models of pain. While some models aim primarily to integrate known and suspected determinants of pain—e.g., the Glasgow (Waddell 1987) and the stress-diathesis (Turk 2002) models—, others may include putative mechanistic details—e.g., associative learning and the resulting fear avoidance behaviors (Flor et al. 2002) or mechanisms underpinning psychological and physiological comorbidities such as depression and stress (Duric and McCarson 2005).

3.5 The requirement of epistemic parity

The key intuition underlying the interventionist argument is that each element of a causal model (each causal claim about a given phenomenon) is equally well supported by an experiment involving controlled interventions. This ensures that no causal factor can be dismissed. Conversely, disparities in terms of empirical support undermine level-neutrality by placing an emphasis on some factors at the expense of others. Thus, epistemic parity among the various elements of a model is a second crucial requirement for level-neutrality. The question that arises now is whether epistemic parity is warranted by the satisfaction of general criteria for evaluating empirical findings, or by the more specific desiderata of interventionism. I offer two arguments in defense of the former, more general view.

First, the strength of the evidence for causal relevance varies depending on how closely real experiments satisfy the requirements for ideal (controlled) interventions. Interventionist accounts of causation assume a sharp distinction between experiments involving ideal or quasi-ideal interventions (experimental interventions, clinical trials) and studies documenting correlations (descriptions of phenomena, imaging studies, observational studies). In practice, however, research may fall somewhere in between these two categories. Different experimental designs provide more or less conclusive evidence for causal relevance depending on how successful they are in eliminating or controlling for confounding variables. For instance, many observational studies do include controls, but fail to demonstrate that controls are identical with the test

condition except for the variable under manipulation. For this reason, randomized controlled trials are considered to be superior to comparative approaches such as cohort and case-control studies (Guyatt et al. 2015).

Likewise, in basic research, some interventions may clearly demonstrate causal relevance to a specific outcome (e.g., gene knockouts), others may lead to a variety of outcomes making it difficult to disentangle between alternate causal interpretations (e.g., gene knockouts resulting in pleiotropic effects), while some experiments are plagued by uncertainties about whether interventions cause the observed outcomes (e.g., neuroimaging studies). At the limit, interventions may not be conducted at all despite overwhelming evidence for strong association (e.g., amyloid deposits and Alzheimer's disease). Yet even if different experimental setups may vary in their ability to demonstrate causation, all of the above constitute scientifically acceptable findings about a given phenomenon and provide evidence for relationships between factors investigated by different disciplines.

My second argument is that interventionism cannot satisfy the requirement of epistemic parity without ultimately relying on the messier assortment of criteria used to evaluate empirical findings in general. Experiments supporting causal claims rely on techniques (instruments, tests, assays, protocols) for measuring specific factors. Different techniques are required for different factors, and the same techniques can be shared by distinct experimental designs aiming to generate evidence for distinct kinds of claims (e.g., causation vs correlation). A distinction must therefore be made between the extent to which a particular experimental design can support claims about causal relationships, and the extent to which a particular technique is suited for measuring a given factor. As discussed above, some experimental designs are better suited for demonstrating causation than others. However, the suitability of a particular design for demonstrating causation doesn't tell us anything about the suitability and quality of the techniques used to measure specific factors. For example, the fact that two experimental designs satisfy interventionist desiderata does not demonstrate that the quality of the evidence for the causal relevance of inflammation to pain is on par with that for the causal relevance of anxiety to pain. The interventionist account doesn't tell us how to compare findings generated by the detection of inflammation using morphological and molecular markers with that generated by psychometric tests assessing anxiety.

The above complications indicate that interventionist desiderata ensure epistemic parity only in conjunction with more fundamental considerations such as reproducibility, reliability, validity and accuracy, which not only apply to empirical findings at large, but also trump the more narrowly focused interventionist criteria.⁸ Woodward acknowledges the more fundamental role of these considerations in establishing the epistemic trustworthiness of various causal claims when he qualifies his statement

⁸ Reproducibility increases the likelihood that relevant correlations and causal contributions can be differentiated from the background noise of chance events and experimental error (Baetu 2013; Bogen and Woodward 1988). When results are well-established, reproducibility can be used as a reliability and accuracy test for calibrating or troubleshooting measurement assays and associated instruments, techniques and experimental setups (Franklin 1986, Chs. 6–7). Validity encompasses evidence that the test is measuring what is supposed to measure and may be established by a systematic cross-referencing of methods and their results (Culp 1994, 1995; Hacking 1983, Ch. 10; Weber 2005, Ch. 9; Wimsatt 1981). Accuracy encompasses sensitivity and specificity, which are measures of the ability to rule out false negatives and false positives.

about mixing lower and higher-level factors by adding “as long as it is true that interventions that change their values are reliably and stably correlated with changes in their putative effects” (2008, 222). Indeed, if the results of a particular study demonstrating a causal relationship cannot be reproduced, then it doesn’t matter that the study satisfies interventionist desiderata (Open Science Collaboration 2015). Consistent lack of reproducibility by studies aiming to replicate the finding using the same setup and techniques is a strong indication that something went wrong in the original study. By comparison, a consistently replicated correlation is a trustworthy finding; if the two findings concern the same phenomenon, researchers will typically include the latter in their models, and exclude the former. Likewise, findings amply cross-referenced by means of different methods and findings obtained by more thoroughly validated and accurate methods are given more epistemic credence than findings not corroborated by alternative methods or findings relying on poorly validated and less accurate methods. This is true whether or not the studies satisfy the conditions for ideal interventions.⁹

3.6 In summary

The emerging picture is that of a nested, triple understanding of the notion of level-neutrality. Even though different disciplines rely on different experimental setups, instruments and techniques required to perform interventions and measure various factors, multidisciplinary models remain level-neutral in as much as they aggregate correlates and causal determinants of a given phenomenon discovered by methods that satisfy the same basic requirements of good experimental practice. This notion of level-neutrality reveals an important way in which empirical research is independent from theoretical frameworks and metaphysical assumptions about levels of explanation and reality associated with particular disciplines.

Causal models and the causal component of multidisciplinary models may further satisfy the more specific desiderata associated with ideal interventions in addition to the basic requirements of experimental practice. This supports the view that there is nothing intrinsic to experimental results that can justify a sharp contrast between intra- and inter-level causality, which is one of the main aims of Woodward’s argument. Finally, in as much as causal models amount to explanations, level-neutrality reveals the non-reductive character of explanations in clinical sciences, which don’t distinguish between somatic and psychogenic etiology/pathology, as argued by Campbell and Kendler.

4 Answers to two objections

4.1 Experimental reductionism

An immediate objection against level-neutrality is that epistemic parity is never understood as strict equality and is certainly not taken to entail that all empirical findings are

⁹ Similar qualifications are made in respect to epidemiological models, most notably in respect to the strength (statistical significance) and consistency (reliability) of correlations (Hill 1965).

equally trustworthy. As pointed out earlier, consistently replicated results are given priority over unreproducible findings, results obtained by means of more thoroughly validated tests trump the results of less well validated tests, and more accurate tests are preferred over less accurate ones. Could it be then that systematic inequalities can create a bias such that results produced by means of thoroughly validated, more reliable or more accurate methodologies associated with some disciplines are likely to be better represented in multidisciplinary models or even trump weaker results obtained by less thoroughly validated, less reliable and accurate methodologies of other disciplines?

For instance, despite the insistence that medical practice should embrace biopsychosocial models, it might be objected that there is a stronger correlation between ‘heart attacks’ and biological factors such as myocardial ischemia, electrical instability and cell damage, than between ‘heart attacks’ and social situations, such as intense grief. Also, the methods for measuring biological factors such as electrical instability are more reliable and enjoy extensive validation as compared to those for measuring grief. Issues can also arise from differences in sensitivity and specificity. For example, let us assume that a social factor, for instance spousal support, involves physical contact, such as a pat on the back. If it turns out that the causal contribution of spousal support is specifically mediated by physical contact irrespective of the social context, then we can dispense altogether of the less specific social-level description and associated methods of assessment, and replace it with the more specific biological-level description and associated experimental methodology. Thus, what is suggested above is that different disciplines and their empirical findings are assigned different levels of epistemic confidence depending on the degree to which they satisfy the basic requirements of experimental research. If these differences are significant enough, we can expect an unequal representation of empirical findings favoring findings generated by some disciplines, which are in turn more likely to be taken into consideration by proposed explanations, thus leading to distinctions in terms of levels of explanation.¹⁰

It seems undeniable that, at least in some situations, this is indeed the case. However, I would like to suggest that there cannot be a complete replacement of the experimental methodologies of one discipline with those of another discipline. Molecular biology never fades into chemistry for the simple reason that molecular biologists are not studying molecules for the sake of studying molecules; they are studying the role of molecules vis-à-vis biological activity (Astbury 1961). It is inconceivable that one can study molecular mechanisms without relying on experimental techniques specific to biology in order to detect and measure the biological phenomena produced by these mechanisms. Interventions on molecular components (gene knockouts, enzyme inhibitors, etc.) must make a difference for biological activity (wing development, immune response, cell growth, etc.).

A similar comment applies to psychological and psychiatric research. For instance, the fact that surgical interventions (Dong et al. 1996) or damage to specific brain areas (Berthier et al. 1988) can induce the dissociation of a sensory-discriminative

¹⁰ I would like to thank Takuya Niikawa for clarifying this very helpful example.

component of pain experience from an affective-motivational component was taken as an indication that the nociceptive input is processed along partially distinct neural pathways. Converse imaging studies of subjects under hypnotic suggestion further supported this hypothesis by demonstrating distinct neural correlates matching surgically-induced dissociation of the two components of pain experience (Rainville et al. 1997). While new methods from neuroscience are introduced in pain research, nothing here indicates that psychological methods of investigation are in any way replaced by biological ones. This should not be particularly surprising, given that researchers must demonstrate the causal relevance of biological structures to psychological activity, in this case changes in overall pain experience or its more specific sensory, affective and cognitive dimensions.

What about cases where psychological diagnosis and its experimental protocols (e.g., self-report, cognitive tasks, behavioral descriptions) were in fact replaced by biological ones (e.g., neurological and physiological correlates, molecular markers)? Various examples come to mind, such as the tracking of eye movements as a proxy for measuring attention, or cortisol levels as a measure of psychological stress. It is tempting to think that because ‘lower-level’ biological forms of assessment cannot be easily faked or distorted, they are preferable to their ‘higher-level’ psychological counterparts. There is nothing wrong with this kind of reasoning. In the case of pain assessment, it is certainly the case that self-reports are susceptible to malingering, as well as to over and under-report, a state of affairs that motivates a legitimate quest for higher specificity forms of assessment less susceptible to manipulation by patients and less open to interpretation by clinicians. Nevertheless, claims about replacement and elimination stem from superficial observations of what clinicians and scientists do, without taking into consideration the scientific justification of these practices. If the ultimate goal is to study psychological phenomena, then physiological tests must be validated against psychological ones. It makes no sense to talk about ‘physiological correlates’ and ‘molecular markers’ if said correlates and markers fail to correlate and act as indicators of subjective experience. Even in cases where biological tests did replace psychological tests in clinical and experimental practice, this replacement was at some point justified by a set of validation procedures whereby the physiological tests were shown to consistently give the same results as a set of psychological tests conducted under rigorously controlled situations typically not achieved in most clinical and experimental setups. In such cases, physiological tests are in fact extensions of earlier psychological tests motivated by the discovery of biological factors correlated with or causally relevant to a psychological phenomenon.¹¹

I propose therefore that, because the methods of assessment associated with what are commonly perceived as ‘higher-level’ disciplines can never be fully substituted by those of ‘lower-level’ disciplines, there is a strong motivation for maintaining a high standard of experimental research across disciplines by developing and perfecting highly reliable, valid and accurate tests.

¹¹ The validation of cortisol secretion patterns as biomarkers of psychological stress (Smyth et al. 1997) illustrates this point.

4.2 Causal exclusion

The second objection I want to address is a version of the causal exclusion argument. A common rationale for maintaining epistemic and metaphysical distinctions between lower and higher levels stems from a widespread physicalist view that higher-level properties or states are determined by lower-level ones, as assumed in supervenience and realization accounts. One implication of physicalism is that if a causal model cites higher-level factors, then the model is expected to expand ‘downwards’ in order to include the lower-level factors determining the higher-level ones. For instance, if certain forms of learning and memory are causally relevant to pain experience, then the underlying cellular mechanisms of long term-potentialiation will also be relevant; and if long-term potentiation is relevant, then so will be the underlying molecular mechanisms regulating the expression of NMDA receptors. Conversely, a model citing lower-level factors may also expand ‘upwards’, in order to include a variety of higher-level factors realized by lower-level factors, which in turn may open new avenues for research and treatment.

While there are experimental limits to how far a model can actually expand, there is an in principle possibility that any given multidisciplinary model will ultimately grow to encompass several layers of causes—for instance, a ‘higher-level’ set of psychological determinants supervening on a ‘lower-level’ set of biological causes—, where each layer seems sufficient to account for the same range of effects on the phenomenon to be explained. If one layer of determinants suffices to cause and explain a given effect on the phenomenon of interest and if we further dismiss the possibility of systematic causal overdetermination, it would seem that any additional causal and explanatory contributions from other layers are redundant. Thus, one may worry that as experimental knowledge accumulates, multidisciplinary models are bound to disintegrate into a multitude of competing single-level explanations.¹²

Many authors dismiss the genuine possibility of multiple explanations of the same explanandum. Jaegwon Kim famously defends a principle of causal exclusion according to which “[n]o event can be given more than one *complete* and *independent* explanation” (1993, 239). In turn, causal exclusion may serve to justify either a reductionism to lower-level causal explanations (Bickle 1998, Ch. 1; Kim 2005, Ch. 4), or antireductionist approaches defending the explanatory autonomy of sciences dealing with higher-level factors (Fodor 1974; Putnam 1975). One way or another, the level-neutrality thesis is undermined.

In response to these worries, I argue that the problem of causal exclusion does not concern multidisciplinary models because they are not complete explanations demonstrating the causal sufficiency of any particular cluster of causal determinants. In his 1993 paper, Kim points out that there can be two “correct explanations only if either at least one of the two is incomplete or one is dependent on the other” (1993, 257). Unfor-

¹² A similar argument can be made in the case of correlations by imagining a case where both the higher and the lower-level factors stand in law-like relationships with respect to the phenomenon of interest. It would seem then that either set of factors is sufficient for deriving equally detailed and accurate descriptions of the phenomenon of interest, thus pointing to two competing deductive-nomological explanations of the same phenomenon.

tunately, he does not provide a clear account of what counts as a causal explanation, and remains silent about the criteria for judging if such an explanation is complete. In a more recent work (2005, Chs. 1–2), he makes the interesting suggestion that causation should be understood in the stronger sense of “generation, or effective production and determination” rather than “mere counterfactual dependence” (2005, 18). While Kim is probably referring to a metaphysical claim about the nature of causation, the distinction has an epistemic counterpart capturing an important difference between causal models and mechanistic explanations. Specifying the nature of this difference can provide a much clearer characterization of the kind of causal explanations prevalent in the life sciences, as well as a clearer understanding of what it means for such explanations to be complete.

Some of the most celebrated scientific explanations in the life sciences amount to descriptions of mechanisms responsible for the production of phenomena (Bechtel 2006, 2008; Craver 2007; Craver and Darden 2013; Darden 2006; Wimsatt 1972). Several characterizations of mechanisms are available, all hinging on the general idea that mechanisms are organized systems of parts causally responsible for producing or maintaining phenomena (Bechtel and Abrahamsen 2005; Glennan 1996, 2002; Illari and Williamson 2011; Machamer et al. 2000). Even though there is a significant overlap between causal models and mechanistic explanations—which should not be surprising given that the former provide the materials and constraints for hypothesizing the latter—the two diverge in several ways, the most obvious one being the shift from causal relevance, which indicates that a phenomenon is susceptible to manipulation via interventions on individual causal determinants or mechanistic components, to the notion of productiveness, which indicates that the mechanistic system as a whole is required to produce a phenomenon (Machamer et al. 2000). Ideally, a mechanistic explanation is complete when it describes a mechanism that is both sufficient to produce and is actually producing the phenomenon in a given experimental setup or biological context (Craver 2006). Sufficiency is understood here in engineering terms, as the ability to construct, physically (e.g., an *in vitro* reconstitution experiment) or *in silico* (e.g., by means of a computer simulation), a system capable of producing the phenomenon of interest starting from parts performing the activities, possessing the properties and being organized as specified in the mechanistic explanation (Baetu 2015, b; Craver and Darden 2013, Ch. 6; Morange 2009; Weber 2005, Ch. 4).

In contrast, causal models don’t provide information about causal sufficiency. That an intervention on determinant X results in a change in phenomenon Y demonstrates causal relevance, but does not prove that X is sufficient to produce Y , or that X is the only determinant of Y . If this is true, then aggregating causally relevant factors into ever more elaborate causal models cannot get us any closer to proving causal sufficiency either (Baetu 2016b). If anything, causal models are bound to remain open-ended, as new causal factors (e.g., new drugs and technologies allowing for novel experimental interventions) can always be appended to the model.

The implication here is that multidisciplinary models fall short of complete mechanistic explanations and therefore cannot conclusively support claims about causal sufficiency. Hence, even if multidisciplinary models expand as postulated by physicalist accounts, this does not automatically entail that such an expansion has the epistemic and metaphysical significance associated with debates about causal

sufficiency, explanatory completeness and their implications for reductionism and antireductionism.¹³

5 Concluding remarks

Using biopsychosocial models of pain as a case study, I defended the level-neutrality of multidisciplinary models. I argued that these models are the epistemic products of empirical research, which neither assumes, nor justifies conceptual or ontological distinctions between levels of description, explanation and reality. The main argument in support of the integration of empirical findings from different disciplines investigating the same phenomenon is grounded in the uniformity of criteria for evaluating empirical claims about correlates, causes and mechanisms. The implication of this thesis is that problems such as inter-level causation, reductionism, and dualism stem from theoretical, rather than empirical considerations. I take this to be a very interesting and significant conclusion, since it suggests that experimental science has the internal resources to promote integrative research which is immune to many of the typical difficulties, objections and debates in contemporary philosophy of science and philosophy of mind.

I am not claiming or suggesting that multidisciplinary models somehow succeed in addressing the explanatory gaps underlying level distinctions. I am very well aware of the fact that researchers face a real theoretical impasse. There is no biopsychosocial theory accounting for the causal contributions of psychological and social factors as ‘fundamental’ constituents in the same way chemical interactions are the basic building blocks of molecular biology. Nor there seems to be any systematic way in which psychological and social factors can be analyzed into finer-grained substrates of a mechanical or chemical sort, which could then receive the same theoretical treatment as mainstream physiological and molecular mechanisms. Given the absence of a unified theoretical framework for conceptualizing biopsychosocial interactions, be it of a holistic or reductive variety, it seems inevitable that pain determinants remain fragmented into biological, psychological and social theoretical kinds as they become part of distinct, yet incomplete biological, psychological and social explanatory perspectives on pain experience. Rather, what I am claiming is that the scientific understanding of psychological phenomena is subjected to an ongoing process of multidisciplinary

¹³ Craver (2007, 223) offers a different answer to the problem of causal exclusion, arguing that “there are generalizations expressing contrastive relations of causal relevance that are true of realized properties and that are not true of their realizers”. Woodward’s (2008, 249–50) response follows a similar line of argumentation, emphasizing that information about causes and effects should be presented in a “parsimonious way” in order to avoid that “certain candidates for causes are too detailed or specific for the effects we want explained” or vice versa. Note, however, that it is not clear how these epistemic virtues fit with the more fundamental requirements for reproducibility, validity and accuracy driving experimental and clinical research. Reproducibility implies a minimal degree of regularity (Baetu 2013), but it is not clear whether this amounts to the kind of generality and parsimony Craver and Woodward have in mind. Validity and accuracy, on the other hand, may and often do favor less general causal dependencies, especially when the issue of external validity arises (Baetu 2016a; Germain and Baetu 2017).

integration driven by experimental findings despite the fact that theoretical frameworks lag behind, remaining fragmented among the disciplines involved.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Ankeny, R. (2001). Model organisms as models: Understanding the lingua franca of the human genome project. *Philosophy of Science*, 68, S251–S261.
- Asmundson, G. J. G., & Wright, K. D. (2004). Biopsychosocial approaches to pain. In T. Hadjistavropoulos & K. D. Craig (Eds.), *Pain: Psychological perspectives*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Astbury, W. (1961). Molecular biology or ultrastructural biology? *Nature*, 190(4781), 1124.
- Baetu, T. M. (2012). Filling in the mechanistic details: Two-variable experiments as tests for constitutive relevance. *European Journal for Philosophy of Science*, 2(3), 337–353.
- Baetu, T. M. (2013). Chance, experimental reproducibility, and mechanistic regularity. *International Studies in History and Philosophy of Science*, 27(3), 255–273.
- Baetu, T. M. (2014). Models and the mosaic of scientific knowledge. The case of immunology. *Studies in History and Philosophy of Biological and Biomedical Sciences*, 45, 49–56.
- Baetu, T. M. (2015a). The completeness of mechanistic explanations. *Philosophy of Science*, 82(5), 775–786.
- Baetu, T. M. (2015). From mechanisms to mathematical models and back to mechanisms: Quantitative mechanistic explanations. In P.-A. Braillard & C. Malaterre (Eds.), *Explanation in biology. An enquiry into the diversity of explanatory patterns in the life sciences* (pp. 345–363). Dordrecht: Springer.
- Baetu, T. M. (2016a). The big picture: The problem of extrapolation in basic research. *British Journal for the Philosophy of Science*, 67(4), 941–964.
- Baetu, T. M. (2016b). From interventions to mechanistic explanations. *Synthese*, 193, 3311–3327.
- Bechtel, W. (2006). *Discovering cell mechanisms: The creation of modern cell biology*. Cambridge: Cambridge University Press.
- Bechtel, W. (2008). *Mental mechanisms: Philosophical perspectives on cognitive neuroscience*. New York: Routledge.
- Bechtel, W., & Abrahamsen, A. (2005). Explanation: A mechanist alternative. *Studies in History and Philosophy of Biological and Biomedical Sciences*, 36, 421–441.
- Bechtel, W., & Richardson, R. (2010). *Discovering complexity: Decomposition and localization as strategies in scientific research*. Cambridge, MA: MIT Press.
- Berthier, M., Starkstein, S., & Leiguarda, R. (1988). Asymbolia for pain: A sensory-limbic disconnection syndrome. *Annals of Neurology*, 24(1), 41–49.
- Bickle, J. (1998). *Psychoneural reduction: The new wave*. Cambridge: MIT Press.
- Bogen, J., & Woodward, J. (1988). Saving the phenomena. *The Philosophical Review*, 97(3), 303–352.
- Bolker, J. (2009). Exemplary and surrogate models: Two modes of representation in biology. *Perspectives in Biology and Medicine*, 52(4), 485–499.
- Broadbent, A. (2013). *Philosophy of epidemiology*. Houndmills: Palgrave Macmillan.
- Campbell, J. (2008). Causation in psychiatry. In K. Kendler & J. Parnas (Eds.), *Philosophical issues in psychiatry* (pp. 196–216). Baltimore: Johns Hopkins University Press.
- Campbell, J. (2013). Causation and mechanisms in psychiatry. In K. W. M. Fulford, M. Davies, G. Graham, J. Sadler, G. Stanghellini, & T. Thornton (Eds.), *Oxford handbook of philosophy and psychiatry* (pp. 935–949). Oxford: Oxford University Press.
- Chalmers, D. (1996). *The conscious mind*. Oxford: Oxford University Press.
- Craig, K. D., & Versloot, J. (2011). Psychosocial perspectives on chronic pain. In M. E. Lynch, K. D. Craig, & P. W. H. Peng (Eds.), *Clinical pain management: A practical guide*. Oxford: Blackwell.
- Craver, C. (2006). When mechanistic models explain. *Synthese*, 153, 355–376.

- Craver, C. (2007). *Explaining the brain: Mechanisms and the mosaic unity of neuroscience*. Oxford: Clarendon Press.
- Craver, C., & Darden, L. (2013). *In search of biological mechanisms: Discoveries across the life sciences*. Chicago, IL: University of Chicago Press.
- Culp, S. (1994). Defending robustness: The bacterial mesosome as a test case. In D. Hull, M. Forbes, & R. Burian (Eds.), *PSA 1994* (pp. 47–57). East Lansing: Philosophy of Science Association.
- Culp, S. (1995). Objectivity in Experimental inquiry: Breaking data-technique circles. *Philosophy of Science*, 62, 430–450.
- Darden, L. (2006). *Reasoning in biological discoveries: Essays on mechanisms, interfield relations, and anomaly resolution*. Cambridge: Cambridge University Press.
- Dennett, D. C. (1996). *Content and consciousness* (2nd ed.). London: Routledge.
- Dong, W. K., Hayashi, T., Roberts, V. J., Fusco, B. M., & Chudler, E. H. (1996). Behavioral outcome of posterior parietal cortex injury in the monkey. *Pain*, 64(3), 579–587.
- Duric, V., & McCarron, K. E. (2005). Hippocampal neurokinin-1 receptor and brain-derived neurotrophic factor gene expression is decreased in rat models of pain and stress. *Neuroscience*, 133, 999–1006.
- Engel, G. L. (1977). The need for a new medical model: A challenge for biomedicine. *Science*, 196(4286), 129–136.
- Engel, G. L. (1980). The clinical application of the biopsychosocial model. *American Journal of Psychiatry*, 137(5), 535–544.
- Feigl, H. (1967). *The mental and the physical. The essay and a postscript*. Minneapolis: University of Minnesota Press.
- Flor, H., Knost, B., & Birbaumer, N. (2002). The role of operant conditioning in chronic pain: An experimental investigation. *Pain*, 95, 111–118.
- Fodor, J. (1974). Special sciences (or: The disunity of science as a working hypothesis). *Synthese*, 28, 97–115.
- Franklin, A. (1986). *The neglect of experiment*. Cambridge: Cambridge University Press.
- Frith, C. D. (1992). *The cognitive neuropsychology of schizophrenia*. Hove: Lawrence Erlbaum Associates.
- Germain, P.-L., & Baetu, T. M. (2017). Extrapolation in biomedical research: A multi-model perspective. In Marco Nathan & Giovanni Boniolo (Eds.), *Foundational issues in molecular medicine*. New York: Routledge.
- Glennan, S. (1996). Mechanisms and the nature of causation. *Erkenntnis*, 44, 49–71.
- Glennan, S. (2002). Rethinking mechanistic explanation. *Philosophy of Science*, 69, S342–S353.
- Gori, G. B. (1989). Epidemiology and the concept of causation in multifactorial diseases. *Regulatory Toxicology and Pharmacology*, 9(3), 263–272.
- Gossel, P. P. (1992). Need for standard methods: The case of american bacteriology. In A. Clarke & J. Fujimura (Eds.), *The Right tools for the job: At work in twentieth century life sciences*. Princeton: Princeton University Press.
- Guyatt, G., Rennie, D., Meade, M. O., & Cook, D. J. (2015). *Users guides to the medical literature: Essentials of evidence-based clinical practice* (3rd ed.). New York: McGraw-Hill.
- Hacking, I. (1983). *Representing and intervening*. Cambridge: Cambridge University Press.
- Hill, A. B. (1965). The environment and disease: Association or causation? *Proceedings of the royal society of medicine*, 58, 295–300.
- Howick, J. (2011). *The Philosophy of evidence-based medicine*. Oxford: BMJ Books.
- Illari, P., & Williamson, J. (2011). Mechanisms are real and local. In P. McKay Illari, F. Russo, & J. Williamson (Eds.), *Causality in the sciences* (pp. 818–844). Oxford: Oxford University Press.
- Kendall, R. E. (2001). The distinction between mental and physical illness. *The British Journal of Psychiatry*, 178(6), 490–493.
- Kendler, K. S., & Campbell, J. (2009). Interventionist causal models in psychiatry: Repositioning the mind-body problem. *Psychological Medicine*, 39, 881–887.
- Kim, J. (Ed.). (1993). Mechanism, purpose, and explanatory exclusion. In *Supervenience and mind*. Cambridge: Cambridge University Press.
- Kim, J. (2005). *Physicalism or something near enough*. Princeton, NJ: Princeton University Press.
- Krieger, N. (1994). Epidemiology and the web of causation: Has anyone seen the spider? *Social Science and Medicine*, 39(7), 887–903.
- Leonelli, S. (2009). On the locality of data and claims about phenomena. *Philosophy of Science*, 76, 737–749.

- Leonelli, S. (2007). What is in a model? In M. Laubichler & G. Muller (Eds.), *Modeling biology: Structures, behaviours, evolution* (pp. 15–36). Cambridge, MA: MIT Press.
- Levine, J. (1983). Materialism and qualia: The explanatory gap. *Pacific Philosophical Quarterly*, 64, 354–361.
- Machamer, P., Darden, L., & Craver, C. (2000). Thinking about mechanisms. *Philosophy of Science*, 67, 1–25.
- McAllister, J. W. (1997). Phenomena and patterns in data sets. *Erkenntnis*, 47, 217–228.
- Melzack, R., & Wall, P. D. (1965). Pain mechanisms: A new theory. *Science*, 150, 971–979.
- Morange, M. (2009). Synthetic biology: A bridge between functional and evolutionary biology. *Biological Theory*, 4(4), 368–377.
- Müller-Wille, S. (2007). Hybrids, pure cultures, and pure lines: From nineteenth-century biology to twentieth-century genetics. *Studies in History and Philosophy of Biological and Biomedical Sciences*, 38(4), 796–806.
- Murphy, D. (2006). *Psychiatry in the scientific image*. Cambridge, MA: MIT Press.
- OpenScienceCollaboration. (2015). Estimating the reproducibility of psychological science. *Science*, 349(6251), 943–951.
- Putnam, H. (Ed.). (1975). Philosophy and our mental life. In *Mind, language, and reality*, (pp. 291–303). Cambridge: Cambridge University Press.
- Rainville, P., Duncan, G. H., Price, D. D., Carrier, B., & Bushnell, M. C. (1997). Pain affect encoded in human anterior cingulate but not somatosensory cortex. *Science*, 277(5328), 968–971.
- Rosenthal, D. M. (1994). Identity theories. In S. Guttenplan (Ed.), *A companion to the philosophy of mind* (pp. 348–355). Oxford: Blackwell.
- Rossy, L. A., Buckelew, S. P., Dorr, N., Hagglund, K. J., Thayer, J. F., McIntosh, M. J., et al. (1999). A meta-analysis of fibromyalgia treatment interventions. *Annals of Behavioral Medicine*, 21(2), 180–191.
- Schaffner, K. F. (1993). *Discovery and explanation in biology and medicine*. Chicago: University of Chicago Press.
- Smart, J. J. C. (1959). Sensations and brain processes. *Philosophical Review*, 68, 141–156.
- Smyth, J. M., Ockenfels, M. C., Gorin, A. A., Catley, D., Porter, L. S., Kirschbaum, C., et al. (1997). Individual differences in the diurnal cycle of cortisol. *Psychoneuroendocrinology*, 22(2), 89–105.
- Steel, D. (2007). *Across the boundaries: Extrapolation in biology and social science*. Oxford: Oxford University Press.
- Turk, D. C. (2002). A diathesis-stress model of chronic pain and disability following traumatic injury. *Pain Research and Management*, 7, 9–19.
- Waddell, G. (1987). A new clinical model for the treatment of low back pain. *Spine*, 12, 632–644.
- Weber, M. (2005). *Philosophy of experimental biology*. Cambridge: Cambridge University Press.
- Weber, M. (2008). Experimentation. In S. Sarkar & A. Plutynski (Eds.), *A companion to the philosophy of biology*. Malden, MA: Blackwell.
- Wimsatt, W. C. (1972). Complexity and organization. In K. F. Schaffner & R. S. Cohen (Eds.), *PSA 1972, Proceedings of the philosophy of science association* (pp. 67–86). Dordrecht: Reidel.
- Wimsatt, W. C. (1981). Robustness, reliability, and overdetermination. In M. B. Brewer & B. E. Collins (Eds.), *Scientific inquiry and the social sciences* (pp. 124–163). San Francisco: Jossey-Bass.
- Woodward, J. (2002). What is a mechanism? A counterfactual account. *Philosophy of Science*, 69, S366–S377.
- Woodward, J. (2003). *Making things happen: A theory of causal explanation*. Oxford: Oxford University Press.
- Woodward, J. (2008). Cause and explanation in psychiatry: An interventionist perspective. In K. Kendler & J. Parnas (Eds.), *Philosophical issues in psychiatry: Explanation, phenomenology and nosology*. Baltimore: Johns Hopkins University Press.