European Journal for Philosophy of Science (2019) 9: 17
https://doi.org/10.1007/s13194-018-0238-0

PAPER IN THE PHILOSOPHY OF THE LIFE SCIENCES

Two challenges for a boolean approach to constitutive inference

Jens Harbecke[1]

Received: 27 March 2018 / Accepted: 7 November 2018 / Published online: 12 December 2018

Abstract
This paper discusses two challenges for a Boolean method for establishing constitutive regularity statements which, according to the regularity theory of mechanistic constitution, form the core of any mechanistic explanation in neuroscience. After presenting the regularity definition for the constitution relation and a methodology for constitutive inference, the paper discusses the problem of full variation of tested mechanistic factors and the problem of informational redundancy. A solution is offered for each problem. The first requires some adjustments to the original theory by introducing the technical notion of a set of types satisfying independent instantiability. The second one is resolved by demonstrating that the problem of informational redundancy is based on a confusion that fails to challenge the theory. It is concluded that the methodology of constitutive inference is consistent and plausible with respect to actual practice in neuroscience.

Keywords Mechanistic explanation · Mechanistic constitution · Constitutive inference · Boolean inference method · Multi-level models

1 Introduction

In a series of papers (Harbecke 2010, 2013b, 2014b, 2015a, b), I have defended a regularity theory of mechanistic constitution[1] and a Boolean method for constitutive

[1] See also Couch (2011).

This article belongs to the Topical Collection: EPSA17: Selected papers from the biannual conference in Exeter
Guest Editors: Thomas Reydon, David Teira, Adam Toon

✉ Jens Harbecke
jens.harbecke@uni-wh.de
http://www.jensharbecke.com

1 Philosophy of Science, Department of Economics and Department of Psychology, Witten/Herdecke University, Alfred-Herrhausen-Str. 50, 58448 Witten, Germany

inference. Both are intended as supporting and amending the “mechanistic ideal” of explanation in cognitive science, which has played an important role in the philosophy of cognitive science of the last two decades.

The regularity account essentially claims that the relationship between a higher-level cognitive capacity and its underlying mechanisms is a specific kind of regularity between (conjunctions of) mechanistic types and phenomena types. The Boolean method is used to uncover deterministic dependency relationships among mechanistic factors and phenomena represented by binary variables. The overall aim of this metaphysical and methodological project has been to offer an adequate reconstruction of successful scientific research especially in cognitive science. Moreover, the goal has been to determine the norms of explanation that should be accepted for certain research problems in this field.

In print (cf. Baumgartner and Casini 2017, 223) as well as in discussions at academic conferences, the regularity view of mechanistic constitution and the Boolean methodology for constitutive inference have recently been characterized as failing to meet their declared goals. It has been argued that (i) the methodology suffers from the problem of a full variation of investigated factors. Moreover, it has been suggested that (ii) the methodology fails to do any interesting work as either (a) it already needs to assume what it purportedly yields, or (b), since phenomena are temporally extended, knowing the causal relationships between parts of the mechanism and slices of the phenomenon are all that is needed, or (c) the interventions needed are “fat-handed” and therefore necessarily confounders for a Boolean methodology for constitutive inference.

My aim in this paper is to demonstrate that both general charges lack force against the regularity theory of mechanistic constitution and the associated Boolean method for constitutive inference. With some amendments, my original theory of constitutive inference stands strong with respect to its applicability and informativeness. Objection (i) can be avoided if the notion of a “set of types satisfying independent instantiability” is put to use for the methodology. Objection (ii)(a) is ineffective, as the only case in which the supposed implication holds is the borderline case of full causal knowledge of the mechanism. In all other cases, which means in virtually all cases of actual scientific research, the original method is non-redundant and informative. Criticism (ii)(b) is benign as phenomena investigated in cognitive science rarely satisfy the conditions needed if a theory of causal inference is supposed to identify the mechanistic factors constituting a given phenomenon. Again, the original theory of constitutive inference remains important and informative. Finally, objection (ii)(c) rests on a confusion.

I will proceed as follows. Section 2 offers a brief introduction to the notion of a mechanistic explanation. By reviewing some recent psychological and neuroscientific studies on group conformity, it hints at some reasons why the notion of mechanistic constitution is interesting and important for research in cognitive science. Section 3 summarizes the main ideas of the regularity theory of constitution and the theory of constitutive inference. Section 4 discusses objection (i), or the problem of full variation of types, and it shows how the alleged problem can be solved. Section 5 demonstrates the ineffectiveness of objection (ii) in its three versions (a)-(c). Section 6 eventually points to some actual limitations of the amended theory.

Section 7 summarizes the results and highlights some puzzles in the context of the methodology of constitutive inference that will have to be left for future research.

2 Explanation in neuroscience

An adequate explanation in cognitive science, says the "mechanistic approach", demands the identification, location, and analysis of the mechanisms underlying a to-be-explained phenomenon on several levels (cf. Bechtel and Richardson 1993; Machamer et al. 2000; Craver 2002; Craver 2007). A "mechanism" should be thought of as a set of "...entities and activities organized such that they are productive of regular changes from start or set-up to finish or termination conditions." (Machamer et al. 2000, 3). A slightly different definition has been offered by Bechtel and Abrahamsen (2005) who describe a mechanism as "...a structure performing a function in virtue of its component parts, component operations, and their organization. The orchestrated functioning of the mechanism is responsible for one or more phenomena." (Bechtel and Abrahamsen 2005, 423)

Numerous examples have been offered in the literature to show that the mechanistic ideal of explanation is in fact observed in cognitive science.[2] A popular model often referred to in the debate is the mechanistic explanation of spatial memory acquisition in rats (cf. Churchland and Sejnowski 1992, ch. 5; Craver and Darden 2001, 115-119; Craver 2002, sec. 2; Bickle 2003, chs. 3-5; Craver 2007, 165-170). For an equally prototypical example of a mechanistic explanation of human cognitive capacities that currently receives much attention from various disciplines including economics and political science, consider the research in cognitive science on group conformity in humans.

Group conformity occurs when subjects align their behavior to that of a group or majority as a result of a public pressure to conform. A renown contribution to the topic is the series of studies conducted by Asch (1951, 1956), who investigated behavioral convergence in the so-called "line experiments" revealing the pressure to conform to an erroneous view. Subjects were asked to give their assessment about a comparative length of two lines after a majority of other human agents had expressed an obviously wrong view. A significant number of subjects displayed a willingness to conform to the erroneous view in light of the judgment of the majority. The general phenomenon uncovered in Asch's experiment is believed to occur in many circumstances including political elections and financial markets.

Recent empirical research has revealed that various factors such as empathy, conformity, venturesomeness (Baddeley et al. 2007, 2012), peer recommendation (Chen 2008), and susceptibility to normative influence (Seiler 2012) reinforce group conformity, whereas higher age, gender (=being male), extraversion

[2]Note that there is also a growing body of literature showing that mechanistic explanation is not the only kind of explanation in cognitive science or the life sciences (cf. Rusanen and Lappi 2007; Huneman 2010; Waskan 2011; Rice 2013; Ross 2015).

(Baddeley et al. 2007, 2012), intelligence (Chung et al. 2011) and performance focus (Andersson et al. 2014) inhibit group conforming behavior. The empathetic aspect in group conformity has been associated with an activation of the amygdala complex (Baddeley et al. 2012), which is long known as playing a central role in emotional coding. As De Dreu and Kret (2016) have shown, within the cortico-amygdala circuitry, the oxytocin molecule modulates social emotional functions that permit empathy and trust towards in-group agents (cf. also Carter 2014; Domes et al. 2007). Taken together, these studies suggest the following entities with their activities as central for the cognitive tendency of group conform behavior:[3]

P : The cognitive tendency to conform to group behavior
M_1 : Activation of the empathic neural network / the limbic system
M_2 : Activation of the amygdala complex
M_3. : Activation of oxytocin in the cortico-amygdala circuitry.

The reason why researchers in the field consider types $M_1 - M_3$ as (partially) explaining the cognitive phenomenon of tending towards group conformity ("P") is that they "are the pivotal means in" (De Dreu and Kret 2016, 170), "play a role for" (Baddeley et al. 2012, 24), "are physiological substrates of" (Carter 2014, 18), "are [constitutively] relevant for" (Domes et al. 2007, 1189) or simply "constitute" the phenomenon on various levels. Or in other words, the conjunction of claims stating *constitutive relationships* among these phenomena and mechanisms is what supplies the (partial) explanation according to the understanding of cognitive scientists: The activation of oxytocin in the cortico-amygdala circuitry partially constitutes the activation of the amygdala complex, which in turn partially constitutes the activation of the empathic neural network in humans as a whole. The latter partially (or wholly) constitutes the internal cognitive tendency in humans to conform to an observed group behavior.

The research on group conformity in humans as reconstructed above seems to observe the general norms associated with the mechanistic approach to explanation. However, from a philosophical point of view, an important question arising in the context of such explanations concerns the nature of the constitution relation: What does it mean to say that an active component C_1 constitutes another active component C_2 or a phenomenon? To offer an adequate analysis of the nature of the constitutive relationship, or the relationship of "being the pivotal means in", "playing a role for", "being a physiological substrates of", "being relevant for", and to provide the methods for establishing empirical theories about such relationships, is the aim of the regularity theory of mechanistic constitution. The following section summarizes the main ideas of this theory.[4]

[3]For further important contributions on the brain networks constituting the tendency to group conformity, see Berns et al. (2005) and the study by Klucharev et al. (2009).

[4]Readers familiar with Harbecke (2010), Harbecke (2013b), Harbecke (2014b), Harbecke (2015a), and Harbecke (2015b) and Couch (2011) may want to skip Section 3 and continue with Section 4.

3 Regularity constitution and constitutive inference

At the heart of the regularity account of mechanistic constitution lies the contention that mechanistic components and their activities are non-redundantly sufficient for the phenomenon in question such that their instances are related as parts and wholes. Or in other words, mechanistic constitution is a type-level non-causal deterministic relationship, such that the instantiations of its relata necessarily overlap.

The regularity account has mainly been developed by Couch (2011) and myself (Harbecke 2010, 2013b, 2014b, 2015a, b). We both have argued for the adequacy of the regularity view on the basis of research and modeling examples in cognitive science such as the one presented in Section 2. What researchers typically mean with terms such as "being the pivotal means in", "playing a role for", "being a physiological substrates of", "being relevant for" is not a probabilistic or random relationship. Rather, cognitive scientists treat the cognitive phenomena studied by their models as being unable to occur without the neural mechanism. At the same time, they expect the cognitive phenomenon to be inevitably present as well, once the physiological mechanism is in place.[5] These implicit assumptions povide the basic rationale for the scientists' investigative work. The deterministic understanding of mechanistic constitution is one reason why the regularity theory of mechanistic constitution seems more adequate as a metaphysical account than probabilistic or abductive theories of constitution (cf. Baumgartner and Gebharter 2016; Gebharter 2017b, c; Baumgartner and Casini 2017).

For the definition of regularity mechanistic constitution, I have presupposed that the types figuring in regularities stand for mechanistic types, which are specific kinds of entities realizing a specific activity.[6] The individuals instantiating mechanistic types are taken to be space-time regions. In other words, the regularity theory denies that there is an important ontological distinction to be made between what the mechanists have called "entities and activities" (cf. Machamer et al. 2000, 3). My formal definition uses Greek letters 'ϕ' and 'ψ' as variables for specific mechanistic types. Capital letters '$\mathbf{X}$', '$\mathbf{X}_1$', '$\mathbf{X}_2$',...,'$\mathbf{X}_n$' express conjunctions of types that can be co-instantiated (either in the same individual or in "co-located" individuals). The formulation reads as follows (cf. Harbecke 2010, 275-278; Harbecke 2015a, 329; to improve legibility, type conjunctions such as "$\phi \wedge \mathbf{X}_1$" are always abbreviated to "$\phi\mathbf{X}_1$"):

Constitution : A mechanistic type ϕ constitutes another mechanistic type ψ (written as "$\mathbf{C}\phi\psi$") if, and only if:

(i) ϕ is contained in a minimally sufficient condition $\phi\mathbf{X}_1$ of ψ, such that...
(ii) $\phi\mathbf{X}_1$ is a disjunct in a disjunction $\phi\mathbf{X}_1 \vee \mathbf{X}_2 \vee \ldots \vee \mathbf{X}_n$ of type conjunctions, all of which are minimally sufficient for ψ, such that the disjunction is minimally necessary for ψ, and such that...

[5]Note that this understanding of constitution reflects what in metaphysics and philosophy of mind has been described as a "supervenience relation" (cf. McLaughlin and Bennett 2008).
[6]Compare this approach to the notion of specific variables as proposed in Spohn (2006).

(iii) if ϕ and $\mathbf{X}_1$ are co-instantiated, then (a) their instances are a mereological part[7] of an instance of ψ, and (b) this instance of ψ is a mereological part of the fused instances of ϕ and $\mathbf{X}_1$.

Conditions (i) expresses the sufficiency of a constituting (conjunction of) type(s) for a constituted type or phenomenon. Condition (ii) ensures that the definition allows for alternative constitutive conditions for the same phenomenon. Condition (iii) demands that the phenomenon occurs (at least partially) in the same place and time as the mechanisms that constitute it, and that the phenomenon occupies no less space and time than its complete mechanism. As it stands, **Constitution** defines *partial* constitution primarily since it is about singular types. However, what can be called *complete* constitution is simply a borderline case of partial constitution. It is a condition that involves all types constituting a phenomenon on a given level.

My definition of **Constitution** suggests that the following conjunction of conditionals $\mathbf{M}_{GC}$ is an adequate reconstruction of the explanation of the cognitive tendency to group conforming behavior introduced in Section 2 ('$\mathbf{M}_{GC}$' stands for 'explanatory model of *G*roup *C*onformity'; the second-order operator '$\Rightarrow_c$' is intended to summarize the criteria specified by **Constitution**; below, we will sometimes speak of a type conjunction being "c-minimally sufficient" for another type in this sense; '$\mathbf{Y}_1$', '$\mathbf{Y}_2$', ...,'$\mathbf{Y}_n$' are used to express disjunctions of conjunctions of properties all of which are minimally sufficient for the type on the right-hand side of the conditional):

$$\mathbf{M}_{GC} : \left(M_3\mathbf{X}_3' \vee \mathbf{Y}_3 \Rightarrow_c M_2\right) \wedge \left(M_2\mathbf{X}_2' \vee \mathbf{Y}_2 \Rightarrow_c M_1\right) \wedge (M_1\mathbf{X}_1' \vee \mathbf{Y}_1 \Rightarrow_c P),$$

Proposition $\mathbf{M}_{GC}$ says: "(If an activation of oxytocin in the cortico-amygdala circuitry is instantiated together with certain other properties in an appropriate way, then a specific kind of amygdala complex activation is instantiated in the same place at the same time) and (If this specific kind of amygdala complex activation is instantiated together with certain other properties, then...) and..." and so on. The central contention of the regularity theory of mechanistic constitution is that the regularites thus expressed faithfully capture what cognitive scientists have in mind when they use terms such as "being the pivotal means in", "playing a role for", "being a physiological substrates of", or "being relevant for". The beauty of the approach lies in the fact that it can define mechanistic constitution almost exclusively in extensional terms. The only undefined semantic notion used in the definition is that of spatio-temporal mereology.

If the adequacy of the regularity theory is accepted, the questions remains what conditions and findings are required to establish a hypothesis about a mechanistic constitution relation between two mechanistic types. In (2015b), I have offered a proposal about the inferences necessary and sufficient to this end. It shares some

[7]The mereological theory presupposed here is *General Extensional Mereology* (**GEM**) as explicated by (Varzi 2009).

important ideas with the work on causal inference in discovery in the context of regularity theories of causation (cf. Graßhoff and May 1995; May 1999; Baumgartner and Graßhoff 2004). However, to adapt this methodology to constitutive contexts, several substantive adjustments are required.

Constitutive inferences always depart from three general premises (cf. Harbecke 2015b, 13):

1. The occurrence of a complete constitutive condition, or of a constitutive mechanism, is sufficient for the occurrence of the constituted phenomenon (determination); moreover, if no such complete constitutive condition or mechanism occurs, the phenomenon also does not occur (dependence).
2. The testing situations obey constitutive homogeneity.
3. A constitutive regularity relationship of the form $\mathbf{X} \vee \mathbf{Y} \Rightarrow_c \phi$ and in the sense of **Constitution** is assumed as a hypothesis, where $\mathbf{X}$ stands for a conjunction of types, $\mathbf{Y}$ stands for a disjunction of conjunctions of types, and ϕ stands for a particular phenomenon type whose underlying mechanisms are to be identified.

To illustrate how research is developed before the background of these premises, consider again the mentioned studies on the cognitive tendency to group conformity. When Domes et al. (2007) began their investigation of the role of oxytocin in the amygdala response to emotional faces, the resarchers had to start with an implicit blank hypothesis $\mathbf{M}_{GC}^{M_2}$ (= a part of model $\mathbf{M}_{GC}$ targeting an explanation of type M_2) stating merely that there is at least one constitutive condition of the phenomenon out there (assumption 3.), and this condition contains all mechanistic types (on the respective level) relevant for the phenomenon (cf. assumption 1.):

$$\mathbf{M}_{GC}^{M_2} : \mathbf{X}_3 \vee \mathbf{Y}_3 \Rightarrow_c M_2$$

Domes et al. then compared two groups of subjects that received oxytocin or a placebo intranasally forty-five minutes before the fMRI sessions. The aim was to detect a difference in the subject's relative tendency to react empathetically and trustful towards others. The general design of their series of experiments and their initial finding can be illustrated by data Table 1.

The "yes" and "no" in the second row of the data table indicate that, in the present setting, the amygdalic activitation relevant for empathy and trust occurred only in subjects with higher oxytocin levels. The amygdalic activation was detected through an fMRI image serving as a reliable indicator, or instrumental variable, of the activation. On the basis of this finding, it was concluded that oxytocin partially constitutes

Table 1 Constitutive difference test on the role of oxytocin for empathy

	group 1	group 2
presence of mechanistic test type	M_3	not-M_3
presence of phenomenon?	yes	no

the amygdalic activation.[8] As a result, the researchers were able to transform the blank initial hypothesis $\mathbf{M}_{GC}^{M_2}$ into the informative hypothesis $\mathbf{M}_{GC}^{M_2\prime}$:

$$\mathbf{M}_{GC}^{M_2\prime} : M_3\mathbf{X}_3' \vee \mathbf{Y}_3 \Rightarrow_c M_2$$

$\mathbf{M}_{GC}^{M_2\prime}$ says that, if an activation of oxytocin in the cortico-amygdala circuitry is instantiated together with certain other properties in an appropriate way, then a specific kind of amygdala complex activation is instantiated in the same place at the same time ("in the amygdala").[9]

It is obvious that the inference yielding this conclusion is justifiable only if the two groups of subjects were either completely similar, or at least nearly similar, in all other respects. This prerequisite is reminiscent of Mill's homogeneity condition specified in his "method of difference" for causal relationships (Mill 1882/1843, 280; cf. also hypothesis 2. above: "The testing situations obey constitutive homogeneity."). Because of the practical impossibility to control all factors in any kind of real-life experiment, I have offered a formulation of the homogeneity condition that is just strong enough for constitutive inferences to be decisive (Harbecke 2015b, 14) .

On the basis of simple difference tests such as the one illustrated by data Table 1, it is possible to design more complex difference tests for any finite number of mechanistic types. All that is required is a systematic alteration of the mechanistic factors and a detection of the behavior of the target phenomenon. Depending on the occurrence and non-occurrence of the phenomenon, further mechanistic factors can be identified that are either part of the same, or of a different, minimally sufficient condition of the phenomenon. In other distributions of the occurrence or non-occurrence, no inference is possible. An exemplary analysis of so-called tests-of-four is developed in (Harbecke 2015b, 14-17).

4 Problem (i): systematic variation

The methodology for constitutive inference presented in the previous section faces a practical challenge. Since the inference rules for more complex difference tests with more than 1 investigated factors require complete data tables to be applicable, the apparent problem is that the theory requires the tested mechanistic types to be instantiable in all logically possible combinations. However, precisely this condition often fails for constituents of cognitive and other higher-level phenomena as the constituents are often causally related to one another. A manipulation on a given mechanistic type will often lead to changes in all or many of its effects. A systematic variation of types will fail accordingly. Hence, only in exceptional cases an inference seems possible for constitutive conditions whose elements are causally related.

[8]Note that, this reconstruction simplifies the target data used by Domes et al. was statistical and not binary. Neverthelss, the authors present the results in a binary way by describing the reaction levels as "higher" vs. "lower" and as "enhanced" vs. "normal" (Domes et al. 2007, 1188).

[9]Note that in hypotheses $\mathbf{M}_{GC}^{M_2}$ and $\mathbf{M}_{GC}^{M_2\prime}$, the expressions "$\mathbf{X}_3$" and "$M_3\mathbf{X}_3'$" state exactly the same conditions; the latter merely makes more constituting mechanistic factors explicit than the former.

Figure 1 illustrates a case in which this problem seems to occur. Apart from the type G representing the phenomenon, it involves seven mechanistic types: $\{F_1, F_2, F_3, F_4, F_5, F_6, F_7\}$. The mechanistic types connected by arrows, such as the one leading from F_4 and F_6 into F_7, are connected by causal relationships (= causal minimal sufficiency & necessity). Sometimes single types are causes of other types (as F_3 is by itself causally sufficient&necessary for F_5). In other cases, some types in conjunction form a sufficient and necessary cause for further types (such as F_1 and F_2 with respect to F_4). It is implicitly presupposed that the occurrence or non-occurrence of the phenomenon G can only be determined with the help of an instrumental variable, as it is usually done in cognitive science research (cf. again the study of Domes et al. (2007), which detected the amygdalic activation through an fMRI image). Types F_5 and F_7 are then jointly causally connected to the behavioral variable serving as a proxy for G.

The overall causal structure contains as a part the (highlighted) chain $F_2 \rightarrow F_4 \rightarrow F_7$. The causal connection implies that no situation can be brought about in which F_7 is present but F_2 is not, unless further interventions are deployed in the background. The latter, however, would violate the homogeneity condition, which in turn would preclude any inference to the constitutive relevance of F_2 and F_7 for G. This seems to be a problem in light of the fact that, ex hypothesis, F_2 *does* (partially) constitute G.

Moreover, suppose that the causal chain is "repaired" by an independent intervention bringing about F_7 even when F_2 is inhibited. Then the hypothetical constitutive relationship of F_2 and the phenomenon G cannot be detected, given that G's occurrence or non-occurrence is detectable only via its behavioral type captured by an instrumental variable. The inhibition of F_2 will have no effect on the behavioral type. Again, it becomes impossible to infer the constitutive relevance of F_2 for G despite the fact that the former does in fact constitute the latter.

Fortunately, there is a relatively easy way to solve this potential problem for the methodology. The solution lies in distinguishing those data tables prone to constitutive inferences from those that are not. The methodology is applicable only to data tables listing independently instantiable types. The criterion of independent instantiability can be defined as follows:

Independent Instantiability: A set of mechanistic types $\mathcal{S}$ satisfies the criterion of Independent Instantiability if, and only if, no element of $\mathcal{S}$ is a causal offspring of any other type in $\mathcal{S}$ (in a given context).

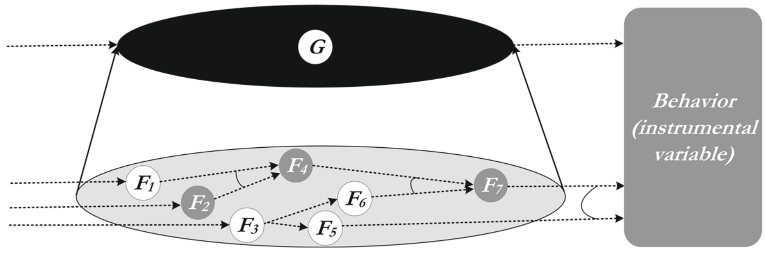

Fig. 1 Mechanistic Causal Chains; dotted arrows represent causal relations; connected dotted arrows represent joint causal sufficiency; lined arrows and ovals represent spatio-temporal containment

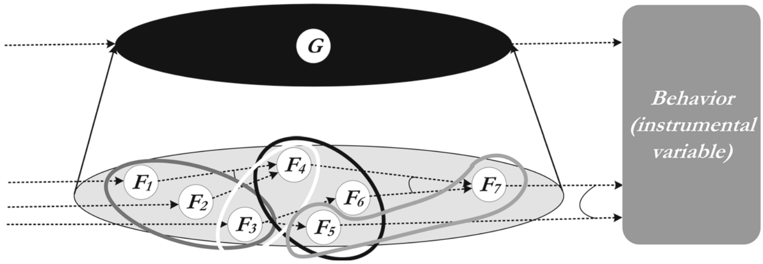

Fig. 2 Mechanistic Slices

The criterion is satisfied by two general situations only: Either (a) the investigated types are part of what can be called a "mechanism slice" of the mechanism underlying the phenomenon, or (b) the types are completely independent and not part of any single causal chain (in the given context). A "mechanism slice" can be defined as follows (it is presupposed here that a set of mechanistic types $\mathcal{S}$ is a mechanism if, and only if, (a) all its elements are connected with another over direct or indirect causal paths and (b) all its elements inhabit the same mechanistic level):

Mechanism Slice: A set of mechanistic types $\mathcal{N}$ is a mechanism slice of a mechanism $\mathcal{M}$ if, and only if, (i) all elements of $\mathcal{N}$ are elements of $\mathcal{M}$, and (ii) no element of $\mathcal{N}$ is a causal offspring of any other type in $\mathcal{N}$.

If we apply the definition of **Independent Instantiability** to the case depicted by Fig. 1, all and only the following sets satisfy the criterion: $\{F_1\}$, $\{F_2\}$, $\{F_4\}$, $\{F_5\}$, $\{F_6\}$, $\{F_7\}$, $\{F_1, F_2\}$, $\{F_1, F_3\}$, $\{F_1, F_5\}$, $\{F_1, F_6\}$, $\{F_3, F_4\}$, $\{F_3, F_7\}$, $\{F_4, F_5\}$, $\{F_4, F_6\}$, $\{F_5, F_6\}$, $\{F_5, F_7\}$, $\{F_1, F_2, F_3\}$, $\{F_1, F_2, F_5\}$, $\{F_1, F_2, F_6\}$, $\{F_4, F_5, F_6\}$, $\{F_1, F_2, F_5, F_6\}$. As it happens, all of these sets are also slices of the mechanism illustrated in the bottom part of the diagram of Fig. 1. Figure 2 picks out some of these slices visually.

It is clear that, once the independent instantiability condition is satisfied by the investigated set of types, the problem of systematic variation is avoided from the start.[10] However, it is important to note that, apart from information about the independent instantiability (in the given context), no previous knowledge about the mechanism and its slices is required for the application of the methodology of constitutive inference.

For illustration, consider again the case depicted by Figs. 1 and 2. After the methodology has identified all sets producible from the set of investigated factors $F_1 - F_7$ that satisfy **Independent Instantiability**, it will investigate the "maximal" sets from this list. Non-maximal sets are sets which can be amended by at least one

[10] One may object that, since F_3 is causally sufficient for F_5 and F_6, any manipulation of F_5 will target F_3 and therefore be a manipulation of F_6 as well. Hence, no systematic variation is possible within the slice. Without going much into depth here, the implicit assumption in the text is that the interventions on types F_5 and F_6 required for the difference test will not have to involve a change in the causes of F_5 and F_6 but change them directly to bring about the situations for comparison. This ensures independent variation within a slice.

more type from the list of investigated types, such that the set thus formed is also on the list. The maximal sets are those sets for which such an amendment is not possible. For the case at hand, all and only the following sets are maximal: $\{F_3, F_4\}$, $\{F_5, F_7\}$, $\{F_1, F_2, F_3\}$, $\{F_4, F_5, F_6\}$, $\{F_1, F_2, F_5, F_6\}$. If run on these sets, the Boolean methodology of constitutive inference will conclude that the following sets of mechanistic types (in conjunction with certain background factors quantified over by $\mathbf{X}_{1-5}$) are minimally sufficient constituting conditions of G:

$$F_1F_2F_3\mathbf{X}_1 \quad F_1F_2F_5F_6\mathbf{X}_2 \quad F_3F_4\mathbf{X}_3 \quad F_4F_5F_6\mathbf{X}_4 \quad F_5F_7\mathbf{X}_5$$

The definition of **Constitution** then implies that the types F_1, F_2, F_3, F_4, F_5, F_6, and F_7 constitute G. A complete mechanistic model will then offer the list of constituting factors of G along with their mutual causal connections. All of this is the desired result. Note that the list of constituting types does not imply that the complex condition $F_1F_2F_3F_4F_5F_6F_7\mathbf{X}'$ is also a minimally sufficient c-condition of G. The simple reason is that it is not minimal.

The question is whether the methodology amended by the independent instantiability requirement will yield the desired result in all cases. The unproblematic cases are those in which some of the investigated types are not part of the mechanism and are not instantiated in the same place and time as the phenomenon. The methodology will clearly exclude these types both because their instantiation does not overlap with the instantiation of the phenomenon, and because their variation will not be accompanied by a detectable change in the phenomenon. In contrast, there are two general kinds of cases which might be expected to cause troubles for the methodology.

The first of these potentially problematic general cases is a mechanistic structure which involves a lower-level causal chain that occurs in the same space-time region as the phenomenon, but that is not in fact constitutive of the phenomenon in question. Figure 3 illustrates such a case. Here the causal chain ... $\rightarrow F_3 \rightarrow F_6 \rightarrow$ ... has no bearing on the behavioral variable and, hence, is constitutively irrelevant for the phenomenon G, whilst the second causal chain ... $\rightarrow F_1F_2 \rightarrow F_4 \rightarrow F_7 \rightarrow$ ... does contain the constitutively relevant factors.

As before, the methodology will now first identify the maximal sets that satisfy the **Independent Instantiability** requirement: $\{F_1, F_2, F_3\}$, $\{F_1, F_2, F_6\}$, $\{F_3, F_4\}$, $\{F_3, F_7\}$, $\{F_6, F_7\}$. The method of difference tests is then applied to each of these sets. As a result, only the following constitutive conditionals will be declared to hold:

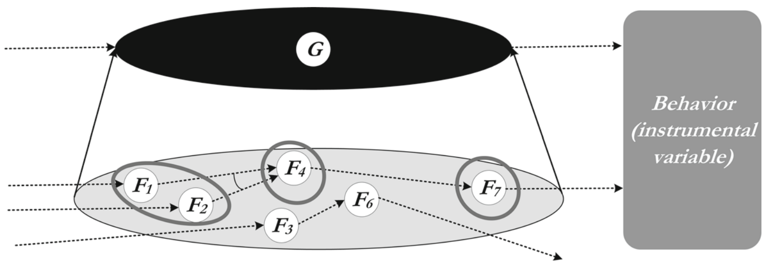

Fig. 3 Irrelevant Mechanisms

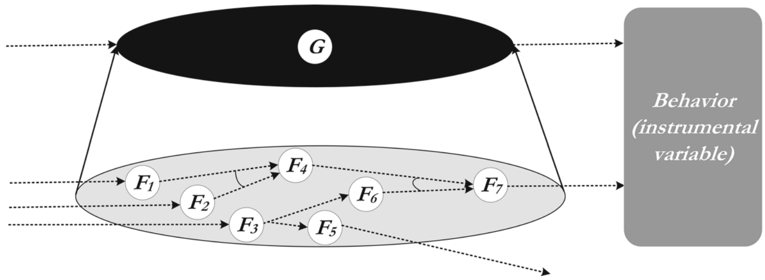

Fig. 4 Mechanistic Dead Ends

$F_1F_2\mathbf{X}_1 \vee \mathbf{Y}_1 \Rightarrow_c G$, $F_4\mathbf{X}_2 \vee \mathbf{Y}_2 \Rightarrow_c G$; $F_7\mathbf{X}_3 \vee \mathbf{Y}_3 \Rightarrow_c G$. No false inference has been made, as neither type F_3 nor type F_6 are declared to constitute G.

The second potentially problematic general case may be a "dead end" scenario. A hypothetical version is illustrated by Fig. 4. In this structure, type F_5 is constitutively irrelevant for G by hypothesis. Again, first the maximal sets satisfying the **Independent Instantiability** requirement must be identified. These are the following ones: $\{F_1, F_2, F_3\}$, $\{F_1, F_2, F_5, F_6\}$, $\{F_3, F_4\}$, $\{F_4, F_5, F_6\}$, $\{F_5, F_7\}$. Again, difference testing is now applied to each set on the list. As a result, the following constitutive relationships will be declared to hold: $F_1F_2F_3\mathbf{X}_1 \vee \mathbf{Y}_1 \Rightarrow_c G$, $F_1F_2F_6\mathbf{X}_2 \vee \mathbf{Y}_2 \Rightarrow_c G$; $F_3F_4\mathbf{X}_3 \vee \mathbf{Y}_3 \Rightarrow_c G$; $F_4F_6\mathbf{X}_4 \vee \mathbf{Y}_4 \Rightarrow_c G$, $F_7\mathbf{X}_5 \vee \mathbf{Y}_5 \Rightarrow_c G$. Also this is as desired. In particular, type F_5 is not listed as a mechanistic type constituting G, and no false inference has been made.

Since the situations illustrated by Figs. 3 and 4 are the only general cases that might lead to a false inference within a Boolean methodology for constitution, the amended methodology can effectively solve the challenge connected to the demand of a full variation of all investigated factors.

5 Problem (ii): analytic redundancy

A second critical remark to the Boolean methodology of constitutive inference that has surfaced in the recent literature (e.g. Baumgartner and Gebharter 2016; Baumgartner and Casini 2017; Gebharter 2017b, c) and in discussions at academic conferences[11] is that the methodology itself does not do any work. The amended theory requires a substantial of causal knowledge about the connections of the investigated mechanistic types before its research strategies can be applied. Moreover, when all causal interconnections among its elements are known beforehand, all constitutive relationships are already known as well. Constitutive difference tests are no longer required in this case. All that seems to be needed is a methodology for causal inference, and there is no further work to be done by the Boolean method for constitutive inference (cf. Gebharter 2017b, ch. 6.4.4).

[11] Lorenzo Casini (University of Geneva) presented this objection at the 3^{rd} Annual Conference Society for the Metaphysics of Science at Fordham University in his "Comment on Harbecke", 5^{th} October 2017.

This critical remark has a true core. It is correct that, if all of the causal connections among the elements of a mechanism $\mathcal{M}$ are known, and if for each element of $\mathcal{M}$, its systematic variation under causal homogeneity is accompanied by a change in the behavioral variable serving as a proxy for the phenomenon in question P, then one can infer that all elements of $\mathcal{M}$ that are instantiated in the same space-time region as P on at least one occasion are constituents of P. Running tests for constitutive inference will then be generally uninformative. However, what the criticism misses is that this implication of the methodology is neither surprising nor does it harm the methodology's original aims and mission.

First, it should be noted that the kind of uninformativeness described by the critic is a direct result of the methodology's implicit assumptions that a) the phenomenon cannot be (or in actual science is usually not) detected directly, but only through its output (= an instrumental type) and b) the phenomenon deterministically ensures the instantiation of its output. If there was a way to directly determine the occurrence or non-occurrence of the phenomenon, constitutive inference would always be informative and would not even have to be restricted to sets satisfying the **Independent Instantiability** requirement. The reason why these implicit assumptions are made is that in the real life of neuroscience and psychology the phenomenon is typically not directly accessible (cf. the comments on Domes et al. 2007 above).

Secondly, in real scientific life, complete causal knowledge of the investigated mechanistic types is virtually never satisfied. As a consequence, in virtually all cases of actual research of cognitive science the Boolean method of constitutive inference can be highly informative. It can identify constitutive relationships even when very limited causal knowledge, namely merely knowledge about causal independencies, is in place. In the borderline case of an investigation of a single factor, no prior causal knowledge is required at all. Secondly, in real scientific life, complete causal knowledge of the investigated mechanistic types is virtually never satisfied. As a consequence, in virtually all cases of actual research of cognitive science the Boolean method of constitutive inference can be highly informative. It can identify constitutive relationships even when very limited causal knowledge, namely merely knowledge about causal independencies, is in place. In the borderline case of an investigation of a single factor, no prior causal knowledge is required at all.

Again, the study by Domes et al. (2007) can serve as an case in point. When the researchers conducted their experiments on the role of oxytocin for social cognition in humans, they knew only a fraction of the causal connections the oxytocin molecule engages in within the brain's physiology. Nevertheless, their experiment yielded highly informative results as they were able to measure the influence of oxytocin administration onto a behavioral variable serving as a proxy for the cognitive phenomenon. Since they only investigated one mechanistic type, the researchers did not have to run a prior test on whether the investigated types are causally independent.

In a similar way, the methodology will yield informative results for the following two general cases.

(A) Nothing is known about a mechanism of a phenomenon ψ, but several types $\phi_1, ..., \phi_n$ are tested for possibly constituting ψ, without it being known whether any of $\phi_1, ..., \phi_n$ is causally connected to the output of ψ.

(B) A causal structure $\mathcal{C}$ is known, and it is asked whether a partial structure $\mathcal{C}'$ contained in $\mathcal{C}$ may be (partially) constitutive of a known phenomenon ϕ, where the output of ψ is not explicitly mentioned in structure $\mathcal{C}$ (possibly, because the output of the microstructure $\mathcal{C}'$ does not have the macroscopic output characteristic of ψ).

In both cases, the methodology of constitutive inference has something informative to say. As before, it first has to identify the maximal sets formed on the basis of the list $\phi_1, ..., \phi_n$, resp. on the basis of the list of types contained in $\mathcal{C}'$, that satisfy the **Independent Instantiability** requirement. It can then run its tests based on the Boolean methodology. No further previous knowledge about the mechanism is required. In particular, neither all of its causal connections among the types, nor their causal connections to the output of the phenomenon have to be known.

Even if the critic agrees, she may still argue that all that is done here is an identification of a causal relation among some mechanistic types and the phenomenon's output. The notion of mechanistic constitution can be completely avoided. After all, the output γ of the phenomenon ψ is the primary key to the occurrence or non-occurrence of the phenomenon. So why can we not just apply the classical methodology for causal inference for the mechanistic types relative to γ?

This objection assumes that, if the variation of the mechanistic constituents of a phenomenon ψ lead to a variation of the output of ψ, then the mechanistic constituents are causes of γ. To see why the assumption is problematic, note that the constituents of ψ often fail to be proportional to γ, in the sense that they are "too microscopic" or "too fine-grained" to be counted among the causes of γ. The constituents of ψ may be proportional to further microscopic or fine-grained types (perhaps those mechanistically constituting γ). However, proportionality may be a key criterion for causal relations, and a lack of match in graininess disqualifies the constituents of ψ as genuine causes of γ, even when a dependency relation among the constituents of ψ and γ can be detected.[12] As a consequence, the classical methodology of causal inference cannot be applied straightforwardly. Thus, it becomes transparent why the Boolean methodology of constitutive inference can be of great analytical value.

A second version of the charge that the methodology itself does not do much work questions the structure of phenomena presupposed by the theory. Assume as most mechanists do that the investigated phenomena are extended in time and space. Then for any constituting factor ϕ of a phenomenon ψ, if you change ϕ from being instantiated ("1") to not being instantiated ("0"), ψ will change its value from 1 to 0 only *after* ϕ's change. Or in other words, the change in ϕ will only cause a change in ψ in a stage ψ_t of ψ occurring slightly later than ϕ. If so, one can just use the known Boolean method of causal inference, and a methodology for constitutive inference is no longer required (cf. Fig. 5).

[12]For authors defending different versions of proportionality, see Yablo 1992; McGrath 1998; Shoemaker 2000; McLaughlin 2007; Schröder 2007; Crane 2008; Woodward 2008; List and Menzies 2009; Weslake 2013; Harbecke 2008, 2013a, 2014a; Harbecke and Atmanspacher 2012; Bernstein 2014.

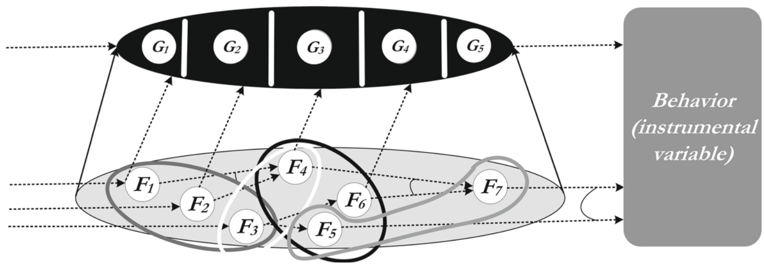

Fig. 5 Temporally sliced phenomena

This version of the objection faces the same charge concerning proportionality as the first one. It is not obvious that the various mechanistic types are good candidates for causes of (slices) of the phenomenon. Further justification is needed.

Another problem is that, for many phenomena studied in cognitive science, the stages envisaged by the objection are not determinable in any intelligible way. As mentioned before, the phenomena investigated in this field are usually measured through instrumental variables only representing effects of the phenomena. Consider again the example of group conformity from Section 2. The cognitive tendency towards group behavior is a cognitive phenomenon whose occurrence or non-occurrence cannot be detected directly, but only via a proxy variable representing the behavior of the subject. There is a principled obstacle for defining stages of this phenomenon inside the subject's head. Only a change of the whole phenomenon is observable, typically through a behavioral variable and perhaps sometimes even directly. It is partially due to this fact about the structure of phenomena studied by cognitive science that a "stages" will not do the job.

A third version of the charge that the methodology itself does not do much work points out that a Boolean inference method faces a problem of systematic confounding through "fat-hand interventions" (Baumgartner and Gebharter 2016, sec. 5; Baumgartner and Casini 2017, 223). The claim is that, whenever a mechanistic type ϕ is manipulated and is accompanied by a change in the phenomenon in question ψ, it is impossible to distinguish between a direct intervention on ψ or an intervention on ψ that went through the indirect path over ϕ. The interventions of types on different levels of mechanistic systems are inevitably "fat-handed". Hence, one can never exclude having ascribed ϕ a constitutive role that it did not have.[13] This criticism confuses a central demand of the methodology.

For illustration, assume a mechanistic structure as the one shown by Fig. 6. In this structure, all mechanistic types in the set $\{F_1, F_2, F_3, F_4, F_6, F_7\}$ are characterized as constitutive of the phenomenon G, whilst the type F_5 is a dead end (cf. also Fig. 4 in Section 4). Now assume we perform an intervention on F_5 (or we compare two situations in which F_5 is instantiated and not instantiated) and a change occurs in G.

[13] Note that the original objection formulated by Baumgartner and Gebharter (2016) and Baumgartner and Casini (2017) differs slightly from the one reconstructed here for my account as the original one focused on problems in Craver's (2007) mutual manipulability account, which adapts a theory of causation by Woodward (2003). Nevertheless, both formulations of the objection share the same fundamental intuition.

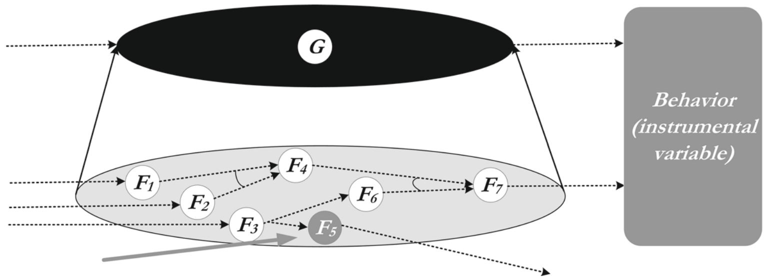

Fig. 6 Intervening on an irrelevant factor

According to our hypothesis, G's change must have come about through a direct path from the intervention onto G as F_5 is a dead end of the mechanism having nothing to do with G and its instrumental variable.

However, if this is the case, then (given the regularity assumption 1. presented in Section 3) also one out of $\{F_1, F_2, F_3, F_4, F_6, F_7\}$ must have changed as illustrated in Fig. 7. In other words, there was a background influence on G through mechanistic types different than F_5. But if so, the homogeneity condition was violated and no constitutive is admissible. Of course, in real-life situations we may be ignorant of such background changes and be thereby led to the false conclusion about the relevance of F_5. However, this does not violate the methodology as such. Its principles still stand as adequate: If all required conditions are satisfied, the methodology will not yield a false conclusion.[14]

6 Actual limitations

The previous sections argued that certain charges against the regularity theory of mechanistic constitution and the Boolean method for constitutive inference lack force. In this section, I want to briefly highlight what I believe are factual limitations of my original account.

The first one concerns the homogeneity condition. The challenge in any actual research situation is, of course, to determine whether homogeneity is in fact satisfied in the sense that in the compared situations or groups only the investigated factor is altered. Or in other words, the knowledge about the satisfaction of homogeneity is necessarily inductive. In any real-life experimental setting, the number of potential confounders is too large to control all of them in a determined way. Whilst this is an undeniable problem for anyone actually doing empirical studies, it is no principled problem for the methodology for constitutive inference. If the condition is satisfied (and the world is as regular as presupposed by virtually any scientific method), then the Boolean method will not lead to false conclusions. In this sense, satisfaction of homogeneity is a practical but not a theoretical problem for the theory.

[14]Note that Gebharter (2017a) has recently promoted a concept of intervention that may mitigate the problem of fat-handedness interventions before it even arises.

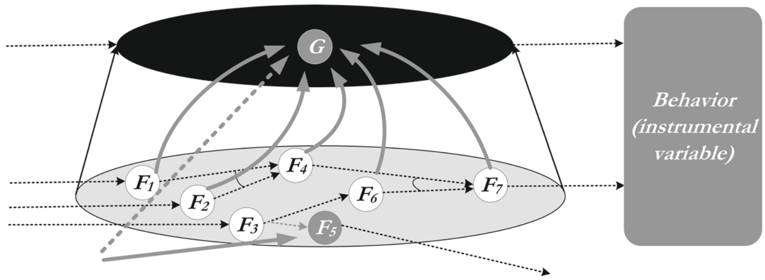

Fig. 7 A violation of homogeneity

A second limitation of the regularity theory of constitution and the methodology for constitutive inference lies in the fact that they have not been supplemented yet with a theory of causation. It was pointed out in Section 4 that a necessary step for the identification of mechanism slices was a test for independent instantiability or causal independence (in a given context). However, no method has been provided that would supply the relevant knowledge. Whilst this is in fact a weakness of the account, there is already helpful literature on a regularity theory of causation and a Boolean method for constitutive inference that could step in to fill this blank (cf. Graßhoff and May 1995; May 1999; Baumgartner and Graßhoff 2004; Baumgartner 2009). In contrast, probabilistic (Suppes 1970; Pearl 2000) or interventionist (Woodward 2003) theories of causation may prove to be not very useful for the Boolean methodology. For instance, if probabilistic causation would be allowed, then conditionalizing on different slices of a mechanism might yield different probabilities for the mechanism's behavior (measured by the instrumental variable).

A third limitation of my regularity-based view of constitution and the method for constitutive inference concerns its capacities for theoretical unification and generalization. As Levy has correctly pointed out, my theory allows "for a division into levels at the local context – within a particular system or mechanism – but not across systems." (2016, 3847) Whilst this may be an obstacle to attain a unity in cognitive science, it does not seem to be a principled problem for the account. It does not imply that the theory goes wrong in its conclusions. The approach merely may not offer good solutions for the unification of models and theories from different fields of cognitive science.

7 Conclusion

The aim of this paper was to answer to two general criticisms of the the regularity view of mechanistic constitution as well as the Boolean methodology for constitutive inference as I have developed them in earlier contributions (Harbecke 2010, 2013b, 2014b, 2015a, b). The first one claimed that (i) the methodology suffers from the problem of a full variation of investigated factors. The second suggested that (ii) the methodology fails to do any interesting work as either (a) it already needs to assume what it purportedly yields, or (b), since phenomena are temporally extended, knowing the causal relationships between parts of the mechanism and slices of the phenomenon are all that is needed, or (c) the interventions needed are "fat-handed" and therefore necessarily confounders for a Boolean methodology for constitutive inference.

I demonstrated that both charges lack force. Objection (i) was solved by introducing the notion of a type set that satisfies an independent instantiability criterion. Objection (ii)(a) was demonstrated to be ineffective, as the only case in which the supposed implication holds is the borderline case of full causal knowledge of the mechanism. In all other cases, and the default situation of actual scientific research, the original method is non-redundant. Criticism (ii)(b) was rejected due to the fact that phenomena investigated in cognitive science rarely satisfy the conditions needed if a theory of causal inference is supposed to identify the mechanistic factors constituting the given phenomenon. Again, the original theory of constitutive inference is important and informative. Finally, objection (ii)(c) was shown to rest on a simple confusion.

Eventually, some actual limitations of the theory were made explicit. These concern the satisfaction of the homogeneity condition, the fact that the method for constitutive inference presupposed a theory of causality that it itself could not provide, and the fact that the theory says little about unification of generalization in cognitive science. Whilst the first problem is unsolvable in principle, the second and third problem are important challenges for future research.

A further open question concerns the applicability of the approach to actual cases of scientific research. Section 2 reviewed an exemplary case that fits well the proposed theory. However, reviewing more cases would be advisable in order to ground the theory more deeply in actual scientific practice and in order to confirm its adequacy.

References

Andersson, M., Hedesström, M., Gärling, T. (2014). A social-psychological perspective on herding in stock markets. *Journal of Behavioral Finance*, *15*(3), 226–234.

Asch, S.E. (1951). Effects of group pressure upon the modification and distortion of judgments. In Guetzkow, H. (Ed.) *Groups, leadership, and men* (pp. 222–236). Pittsburgh: Carnegie.

Asch, S.E. (1956). Studies of independence and conformity: i. a minority of one against a unanimous majority. *Psychological Monographs: General and Applied*, *70*(9), 1.

Baddeley, M., Burke, C., Schultz, W., Tobler, P. (2012). Herding in financial behaviour: a behavioural and neuroeconomic analysis of individual differences. Cambridge Working Papers in Economics. Faculty of Economics, University of Cambridge.

Baddeley, M., Pillas, D., Christopoulos, Y., Schultz, W., Tobler, P. (2007). *Herding and social pressure in trading tasks: a behavioural analysis. Faculty of Economics Cambridge Working Papers in Economics CWPE no. 0730*. Cambridge: University of Cambridge.

Baumgartner, M. (2009). Uncovering deterministic causal structures: a Boolean approach. *Synthese*, *170*(1), 71–96.

Baumgartner, M., & Casini, L. (2017). An abductive theory of constitution. *Philosophy of Science*, *84*(2), 214–233.

Baumgartner, M., & Gebharter, A. (2016). Constitutive relevance, mutual manipulability, and fat-handedness. *The British Journal for the Philosophy of Science*, *67*(3), 731–756.

Baumgartner, M., & Graßhoff, G. (2004). *Kausalität und kausales Schliessen: eine einführung mit interaktiven Übungen*. Bern: Bern Studies in the History and Philosophy of Science.

Bechtel, W., & Abrahamsen, A. (2005). Explanation: a mechanist alternative. *Studies in History and Philosophy of Biological and Biomedical Sciences*, *36*(2), 421–441.

Bechtel, W., & Richardson, R. (1993). *Discovering complexity: decomposition and localization as scientific research strategies*. New York: Princeton University Press.

Berns, G.S., Chappelow, J., Zink, C.F., Pagnoni, G., Martin-Skurski, M.E., Richards, J. (2005). Neurobiological correlates of social conformity and independence during mental rotation. *Biological Psychiatry*, *58*(3), 245–253.

Bernstein, S. (2014). Two problems for proportionality about omissions. *Dialectica*, *68*(3), 429–441.

Bickle, J. (2003). *Philosophy and neuroscience: a ruthlessly reductive account*. Dordrecht: Kluwer.

Carter, C.S. (2014). Oxytocin pathways and the evolution of human behavior. *Annual Review Of Psychology*, *65*, 17–39.

Chen, Y.-F. (2008). Herd behavior in purchasing books online. *Computers in Human Behavior*, *24*(5), 1977–1992.

Chung, D., Yun, K., Kim, J.H., Jang, B., Jeong, J. (2011). Different gain/loss sensitivity and social adaptation ability in gifted adolescents during a public goods game. *PloS One*, *6*(2), e17044.

Churchland, P.S., & Sejnowski, T.J. (1992). *The computational brain*. Cambridge: MIT press.

Couch, M. (2011). Mechanisms and constitutive relevance. *Synthese*, *183*(3), 375–388.

Crane, T. (2008). Causation and determinable properties: on the efficacy of colour, shape, and size. In J. Hohwy, & J. Kallestrup (Eds.)*, Being reduced: new essays on reduction, explanation, and causation* (pp. 176–195). New York: Oxford University Press.

Craver, C. (2002). Interlevel experiments and multilevel mechanisms in the neuroscience of memory. *Philosophy of Science*, *69*(3), 83–97.

Craver, C. (2007). *Explaining the brain*. New York: Oxford University Press.

Craver, C., & Darden, L. (2001). Discovering mechanisms in neurobiology. In Machamer, P., Grush, R., McLaughlin, P. (Eds.) *Theory and method in the neurosciences* (pp. 112–137). Pittsburgh: University of Pittsburgh Press.

De Dreu, C.K., & Kret, M.E. (2016). Oxytocin conditions intergroup relations through upregulated in-group empathy, cooperation, conformity, and defense. *Biological Psychiatry*, *79*(3), 165–173.

Domes, G., Heinrichs, M., Gläscher, J., Büchel, C., Braus, D.F., Herpertz, S.C. (2007). Oxytocin attenuates amygdala responses to emotional faces regardless of valence. *Biological Psychiatry*, *62*(10), 1187–1190.

Gebharter, A. (2017a). Causal exclusion and causal bayes nets. *Philosophy and Phenomenological Research*, *95*(2), 353–375.

Gebharter, A. (2017b). *Causal nets, interventionism, and mechanisms: Philosophical foundations and applications Synthese Library 381*. Dordrecht: Springer.

Gebharter, A. (2017c). Uncovering constitutive relevance relations in mechanisms. *Philosophical Studies*, *174*(11), 2645–2666.

Graßhoff, G., & May, M. (1995). Methodische analyse wissenschaftlichen entdeckens. *Kognitionswissenschaft*, *5*, 51–67.

Harbecke, J. (2008). Mental causation. Investigating the mind's powers in a natural world. Frankfurt a M.: Ontos.

Harbecke, J. (2010). Mechanistic constitution in neurobiological explanations. *International Studies in the Philosophy of Science*, *24*(3), 267–285.

Harbecke, J. (2013a). On the distinction between cause-cause exclusion and cause-supervenience exclusion. *Philosophical Papers*, *42*(2), 209–238.

Harbecke, J. (2013b). Regularity theories of mechanistic constitution in comparison. In Hoeltje, M., Spitzley, T., Spohn, W. (Eds.) *GAP 8 Was duerfen wir glauben? Was sollen wir tun?* (pp. 126–134). Duisburg-Essen: DuEPublico.

Harbecke, J. (2014a). Counterfactual causation and mental causation. *Philosophia*, *42*(2), 363–385.

Harbecke, J. (2014b). The role of supervenience and constitution in neuroscientific research. *Synthese*, *191*(5), 725–743.

Harbecke, J. (2015a). Regularity constitution and the location of levels. *Foundations of Science*, *20*(3), 323–338.

Harbecke, J. (2015b). The regularity theory of mechanistic constitution and a methodology for constitutive inference. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, *54*, 10–19.

Harbecke, J., & Atmanspacher, H. (2012). Horizontal and vertical determination of mental and neural states. *Journal of Theoretical and Philosophical Psychology*, *32*(3), 161–179.

Huneman, P. (2010). Topological explanations and robustness in biological sciences. *Synthese*, *177*(2), 1–33.

Klucharev, V., Hytönen, K., Rijpkema, M., Smidts, A., Fernández, G. (2009). Reinforcement learning signal predicts social conformity. *Neuron*, *61*(1), 140–151.

Levy, A. (2016). The unity of neuroscience: a flat view. *Synthese*, *193*(12), 3843–3863.

List, C., & Menzies, P. (2009). Non-reductive physicalism and the limits of the exclusion principle. *Journal of Philosophy*, *106*, 475–502. http://eprints.lse.ac.uk/20118/.

Machamer, P., Darden, L., Craver, C. (2000). Thinking about mechanisms. *Philosophy of Science*, *67*(1), 1–25.

May, M. (1999). kausales Schliessen. Eine Untersuchung über kausale erklärungen und Theorienbildung. Ph.D. thesis, University of Hamburg, Germany.

McGrath, M. (1998). Proportionality and mental causation: a fit? *Noû,s*, *32*, 167–176.

McLaughlin, B., & Bennett, K. (2008). Supervenience. In Zalta, E.N. (Ed.) *The Stanford Encyclopedia of Philosophy (Fall 2008 ed.)*

McLaughlin, B.P. (2007). Mental causation and shoemaker-realization. *Erkenntnis*, *67*(2), 149–172.

Mill, J.S. (1882/1843). *A system of logic: ratiocinative and inductive, book III*, 8th edn. London: Harper and Brothers.

Pearl, J. (2000). *Causality: models, reasoning and inference*. Cambridge: Cambridge University Press.

Rice, C. (2013). Moving beyond causes: optimality models and scientific explanation. Nous.

Ross, L.N. (2015). Dynamical models and explanation in neuroscience. *Philosophy of Science*, *82*(1), 32–54.

Rusanen, A.-M., & Lappi, O. (2007). The limits of mechanistic explanation in neurocognitive sciences. In Vosniadou, D.K.S., & Protopapas, A. (Eds.) *Proceedings of the European Cognitive Science Conference. Lawrence Erlbaum Associates*.

Schröder, J. (2007). Mental causation and the supervenience argument. *Erkenntnis*, *67*(2), 221–237.

Seiler, M.J. (2012). Forward and falsely induced reverse information cascades. *Journal of Behavioral Finance*, *13*(3), 226–240.

Shoemaker, S. (2000). Realization and mental causation. In *The Proceedings of the 20th World Congress of Philosophy*, (Vol. 9 pp. 23-33).

Spohn, W. (2006). Causation: an alternative. *The British journal for the philosophy of science*, *57*(1), 93–119.

Suppes, P. (1970). *A probabilistic theory of causality*. Amsterdam: North Holland.

Varzi, A. (2009). Mereology. In Zalta, E.N. (Ed.) *Stanford Encyclopedia of Philosophy (Summer 2009 ed.)* http://plato.stanford.edu/entries/mereology/.

Waskan, J. (2011). Mechanistic explanation at the limit. *Synthese*, *183*(3), 389–408.

Weslake, B. (2013). Proportionality, contrast and explanation. *Australasian Journal of Philosophy*, *91*(4), 785–797.

Woodward, J. (2003). *Making things happen: a theory of causal explanation*. New York: Oxford University Press.

Woodward, J. (2008). Mental causation and neural mechanisms. In J. Hohwy, & J. Kallestrup (Eds.)*, Being reduced: new essays on reduction, explanation, and causation* (pp. 218–262). New York: Oxford University Press.

Yablo, S. (1992). Mental causation. *The Philosophical Review*, *101*(2), 245–280.