Causal emergence from effective information: Neither causal nor emergent?

Joe Dewhurst

Munich Center for Mathematical Philosophy, Ludwig-Maximilians-Universität München, Munich, Germany

Correspondence
Joe Dewhurst, Munich Center for Mathematical Philosophy, Ludwig-Maximilians-Universität München, Munich, Germany.
Email: joseph.e.dewhurst@gmail.com

Abstract
The past few years have seen several novel information-theoretic measures of causal emergence developed within the scientific community. In this paper I will introduce one such measure, called ‘effective information’, and describe how it is used to argue for causal emergence. In brief, the idea is that certain kinds of complex system are structured such that an intervention characterised at the macro-level will be more informative than one characterised at the micro-level, and that this constitutes a form of causal emergence. Having introduced this proposal, I will then assess the extent to which it is genuinely ‘causal’ and/or ‘emergent’, and argue that it supports only an epistemic form of causal emergence that is not as exciting as it first seems.

KEYWORDS
causation, complexity, emergence, information theory, intervention, science

1 | INTRODUCTION

The aim of this paper is to provide some philosophical clarity about the nature and scope of a recent formal approach to causal emergence developed by Erik Hoel and colleagues. Doing so is important for two reasons: firstly, because this novel approach has sparked considerable interest and excitement in the scientific communities that Hoel works in, and secondly, because if successful it would offer a unique solution to several outstanding questions concerning reduction...
and emergence in metaphysics, philosophy of science, and philosophy of mind. The approach itself is not necessarily very philosophically innovative, following in the interventionist tradition of Pearl and Woodward, but the new formal details it provides have attracted significant attention outside of philosophy, and so it is important to clarify any potential philosophical confusions that it might encourage. It would also be worthwhile for philosophers themselves to engage with the excitement about this approach, in order to better understand how causation and emergence are conceived of in contemporary scientific circles. Finally, if the approach can be made to work, then it might be able to cast light on other philosophical debates beyond those concerned strictly with causation and emergence, such as the (dis)unity of the sciences and the relationship between different scales of scientific explanation, for example by providing an independent measure of the relative autonomy of one explanatory scale from another. Before it can do any of this, however, we must first assess what the approach is really measuring.

In section 1, I will present Hoel’s new approach and (very briefly) its background in the integrated information theory (IIT) of consciousness. In section 2, I will ask whether the kind of emergence it demonstrates is really causal, and argue that it is only causal in a deflationary sense that it is unlikely to resolve any of the more substantial metaphysical concerns to do with reduction and emergence. In section 3, I will ask whether causation, even in this deflated sense, can be considered emergent under Hoel’s approach, and argue that he has only demonstrated weak or epistemic emergence. Nonetheless, his framework still offers a useful analysis of emergence in this weak sense, and could prove valuable if it is applied with some philosophical caution.

2 EFFECTIVE INFORMATION AND CAUSAL EMERGENCE

The notion of ‘effective information’ was first introduced by Tononi and Sporns (2003) in the context of IIT, a formal approach to the study of consciousness previously developed by Tononi. It is worth briefly saying a little more about IIT before proceeding, as it provides some important background context for Hoel’s model of causal emergence. According to IIT, consciousness is equivalent to some causal property of a system, and can be measured in terms of integrated information (or $\phi$), which is “the amount of information generated by a complex of elements, above and beyond the information generated by its parts” (Tononi, 2008: p. 216). An information-theoretic measure of causal structure therefore plays a central role within IIT. If this measure could also be shown to generate causal emergence, as Hoel aims to do, then IIT could serve as the basis for a non-reductive theory of mind and consciousness, which would of course be very exciting for both philosophers and cognitive scientists. This connection is important when it comes to the potential impact of his approach, but Hoel’s model of causal emergence does not strictly depend upon the success of IIT, and could be interesting or useful in its own right even if IIT were to fail. Furthermore, it has been suggested that IIT should be understood as supporting a hylomorphic (structure-first) ontology (see Owen, 2019), in which case many of the usual arguments against causal emergence might be misdirected. For the remainder of this paper, therefore, I will assess Hoel’s proposal on its own terms, without discussing IIT or hylomorphism directly.

In their 2003 paper, Tononi and Sporns describe effective information (EI) as “a quantity capturing all causal interactions that can occur between two parts of a system” (Tononi & Sporns, 2003, emphasis in original). It is formalised as the mutual information (MI) between two complementary partitions (A and B) of a system (S), where A has been intervened on such that its outputs are maximally entropic (using an independent noise source), to give a measure, EI(A → B), of the effect of this intervention on B. They explicitly state that this measure should be interpreted causally:
Since A is substituted by independent noise sources, the entropy that B shares with A is due to causal effects of A on B [...] Thus, if the connections between A and B are strong and specialised, different outputs from A will produce different firing patterns in B, and EI(A → B) will be high. On the other hand, if the connections between A and B are such that different outputs from A produce scant effects, or if the effect is always the same, then EI(A → B) will be low or zero. (Tononi & Sporns, 2003, emphasis added)

EI measures the effect that an intervention on one part of a system has on another part of that system, and is intended as an information-theoretic measure of the degree to which two parts of a system are causally integrated. A very similar account of causation is given by Griffiths et al. (2015), who define an information-theoretic measure of ‘causal specificity’ that is intended to adjudicate between different possible sources of phenotypic variation in biology (e.g., genetic vs. environmental). Both accounts are descendants of the information-theoretic analyses of causation given by Pearl (1988, 2000; see also Spirtes et al., 1993), which provided some of the inspiration for Woodward’s (2003) interventionist account of causation. Hoel draws this connection explicitly, and I will return to in the next section, as it sheds some important light on exactly what kind of causal emergence his model gets us. For now I will move on to Hoel’s application of EI to the question of causal emergence.

Hoel (2017; see also Hoel et al., 2013; Albantakis et al., 2019) demonstrates how adjusting the granularity of our description of a system can vary the quantity of EI that holds between one state of the system and another, such that there is potentially greater EI (and thus ‘more’ causation) when the system is analysed at the macro-level. He first defines a variant formalisation of EI, where a uniform intervention is applied across all states of the system, and then the results of this intervention are observed. The more certain the effects of an intervention, the greater the EI of the system. This can be most clearly demonstrated in terms of transition probability matrices (TPMs), which show the probabilities of one state transitioning to some other state. Consider three different four-state TPMs (adapted from Hoel, 2017):

\[
\begin{bmatrix}
0 & 0 & 1 & 0 \\
1 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 \\
0 & 1 & 0 & 0
\end{bmatrix}
\text{TPM1}
\]

\[
\begin{bmatrix}
\frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 \\
\frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 \\
0 & 0 & 0 & 1 \\
0 & 0 & 0 & 1
\end{bmatrix}
\text{TPM2}
\]

\[
\begin{bmatrix}
\frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\
\frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\
\frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\
\frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4}
\end{bmatrix}
\text{TPM3}
\]
TPM 1 has high (in fact, maximal) EI: if we know which initial state it is in, then we also know, with certainty, which state it will transition to (here EI = 1). The EI of TPM 2 and TPM 3 is lower: for TPM 2 only states 3 and 4 give us certain predictions, while states 1 and 2 are less predictive (the EI of the whole system = 0.5), whereas TPM 3 is entirely random (EI = 0). Taking EI as a measure of causation, we can say that the states of TPM 1 are more causally integrated than those of either TPM 2 or TPM 3 (the latter of which is not at all causally integrated). This should come as no surprise, as it is fairly intuitive that a stronger causal connection should be more informative about the future state of a system.

What does this have to do with causal emergence? Again, Hoel gives a clear demonstration by comparing the EI of two more TPMs, where this time one is a macroscale description of the other:

\[
\begin{bmatrix}
\frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & 0 \\
\frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & 0 \\
\frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & 0 \\
\frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & 0 \\
\frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & 0 \\
\frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}
\text{TPM4}
\]

\[
\begin{bmatrix}
1 & 0 \\
0 & 1
\end{bmatrix}
\text{TPM5}
\]

In the first, 8-state description (TPM 4), the EI is relatively low (EI = 0.18), as only one state will transition with high probability, while each of the other seven states will transition with equal probability to any of the same seven. This means that if we apply an arbitrary intervention to the system, we would have (on average) relatively little knowledge of which state it will transition to next. Compare this with the second, 2-state description (TPM 5), where the seven unpredictable (micro) states of the first system have been grouped together into a single, now highly predictable, (macro) state. Under this new (macroscale) description, the EI of the system has risen to 1, as any arbitrary intervention will give us perfect knowledge of which state it will transition to next. If, as Hoel claims, EI is an adequate measure of causation, then there is ‘more’ causation in the macroscale description of this system (TPM 5) than in the microscale description (TPM 4), demonstrating a novel kind of causal emergence. Causation emerges under the macroscale description because we can more reliably predict the outcomes of our macroscale interventions, or to put it another way, the kinds of system that can be coarse-grained in the way described above are those that (according to Hoel) exhibit causal emergence.

Rosas et al. (2020, see also Seth, 2010) have recently presented another information-theoretic account of causal emergence in multivariate systems, albeit based on Granger rather than Pearl causality. This is an important difference, and one that deserves further consideration, but I will focus here just on Hoel’s approach. Flack (2017) has also presented an account of downward causation based on coarse-graining that seems compatible with Hoel’s approach, and applied it to various real-world systems including macaque dominance relationships,
protein–protein interactions in molecular biology, and deep neural networks. All three approaches (Rosas et al.’s, Flack’s, and Hoel’s) bear some resemblance to earlier work by James Crutchfield and colleagues on ‘computational mechanics’, see for example Crutchfield (1994a, 1994b), but again I will not be able to consider this antecedent approach in any detail here.

The potential applications of Hoel’s approach should be apparent. In particular, if mental states can be treated as macroscale coarse-grains of microscale neural structures, then this could offer a way of formally demonstrating that the macroscale possesses novel causal powers, thus blocking reductionism or epiphenomenalism of the kind threatened by causal exclusion arguments (cf. Kim, 2000, 2005). However, the sense in which EI measures causation, and the kind of emergence that this measure supports, is not yet clear. I will address each point in turn.

3 | IS IT REALLY CAUSAL?

Hoel’s commitment to an interventionist account of causation is relatively clear. His method for measuring EI is inspired by Pearl (2000), and the sense in which it tracks causal structure is very much interventionist in flavour (cf. Woodward, 2003). Essentially, what EI measures is the extent to which an arbitrary intervention is predictive of the future state of a system, such that a highly structured (or determinate) system will have high EI, and a very chaotic (or indeterminate) system will have low EI.3 Another way of putting this is to say that we can intervene more reliably on a system with high EI, and it is in this sense that EI is a measure of causation.

When we coarse-grain over a system we are focusing our attention on particular aspects of its structure, by grouping together states whose (somewhat chaotic) outcomes we are happy to treat as equivalent (for whatever reason). In doing so we raise the EI of the system, because now our interventions will more reliably bring about the state that we are interested in—a macro-state that is a coarse-grained aggregate of many different fine-grained micro-states (compare TPM 5 with TPM 4). It is only in this interventionist sense that there is ‘more’ causation in the macroscale system than in its microscale equivalent, because we are able to more reliably predict the outcome of an arbitrary intervention (relative to our coarse-grained grouping of the micro-states).

Gillett (2019) distinguishes between “thick” and “thin” senses of causation, with the latter corresponding to the interventionist sense seemingly favoured by Hoel, while the former implies something more metaphysically robust. He argues that neither sense can support anything like emergent downward causation, and while my focus here is rather on the prospects for emergent horizontal causation, his distinction is still a helpful one. Those who tend to get excited about causal emergence are typically interested in something more like emergence in Gillett’s thick sense, that is, the emergence of genuinely novel causal powers, such that (for example) mental states could have causal powers over and above those of their physical supervenience base (see e.g., Robb, 2019). It is not obvious that the thin, interventionist sense of causation used by Hoel gets us anything like causal powers in this sense, rather than an account of causal explanation. Reutlinger (2012) has offered some criticism of interventionism as a fundamental account of causation, and similar concerns are raised by Strevens (2007, 2008). One problem is that an ‘intervention’ is itself already a causal notion, suggesting a degree of circularity if interventionism is understood as a metaphysical account of causation as such, rather than an account of causal explanation. Strevens suggests that it might be better understood as an account of our causal concepts, and as Reutlinger (2012: p. 788) notes, Woodward sometimes seems to agree. In his
response to Strevens' review of *Making Things Happen*, Woodward himself is very clear that he has no metaphysical ambitions, but is rather concerned with questions about the role of causation in scientific methodology (Woodward, 2008: p. 194; cf. Woodward, 2015a).

The interventionist account of causation is distinctively deflationary in flavour. It does not seem like the interventionist approach is well-equipped to tell us whether there are genuinely novel causal powers at the macro-level of some system, rather than merely novel causal explanations. Hoel does seem aware of this distinction, writing that his account “does not contradict other theories of emergence, such as proposals that truly novel laws or properties may come into being at higher scales” (2017: p. 13), so it is possible that he would accept that he has only so far offered an account of emergent causal explanations—but if so, the implications of his approach are distinctively less exciting than he has made them out to be. It would not support the kind of ontological emergence necessary for an anti-reductionist theory of mind or consciousness, but rather just the epistemic emergence of a novel kind of explanatory level, based on the complexity of a system.

A defender of interventionism might object here that interventionist causation is just as ‘real’ as any other kind of causation (see e.g., Andersen, 2017; Braddon-Mitchell, 2017 for recent defences of interventionism as a full-blown account of causation). I am willing to concede this point for the sake of argument, but even if we were to grant that interventionist causation is genuinely causal, it would still be unclear whether the macroscale interventions described by Hoel actually give rise to any ‘more’ causation than there already is at the microscale. Of course, this assumes that we are taking micro-causation as our starting point, but insofar as Hoel's account is an account of causal emergence, it seems fair to ask whether it can demonstrate the emergence of macro-causation over-and-above micro-causation. If it is just intended as an account of causation at the macro-level, rather than an account of causal emergence per se, then it would be much less exciting than it first appears. An information-theoretic account of macro-level causation could be interesting in its own right, but without some foundation in micro-level causation, it would not qualify as an account of causal emergence.

One potential concern here has to do with so-called ‘fat-handedness’ objections (cf. Baumgartner & Gebharter, 2016). If every macroscale structure supervenes upon a microscale structure, then any intervention on the macroscale will at the same time also be an intervention on the microscale, making it hard (or maybe even impossible) to distinguish causation at different scales. Hoel explicitly states that “The set of all possible causal [macro-]models, \( S \), is entirely fixed by the base \( S_m \) [the micro-structure]” (2017: p. 5). He goes on to note that “In technical terms this known as supervenience: given the lowest scale of any system (the base), all the subsequent macro causal models of that system are fixed” (ibid.). His commitment to the supervenience of the (supposedly causally emergent) macro-structure on some underlying micro-structure should therefore not be in question. In order to measure EI at the macro-level, we need to intervene at the macro-level, but doing so will necessarily also involve a micro-level intervention. So even if we can intervene more reliably (at least relative to macro-level outcomes) by targeting the macro-level, it does not seem like there are really any novel causal processes taking place at this level, because corresponding causal interventions and processes will always simultaneously be taking place at the micro-level. This is not to say that there cannot be macro-level causation in some sense (cf. Shapiro & Sober, 2007; Woodward, 2015b), but rather that if EI were measuring an increase in causal power at the macro-level, then there would also need to be some commensurate change at the micro-level, which after all has just undergone the very same intervention (in the sense that any intervention must simultaneously target both macro- and micro-level).
None of this is to say that Hoel's argument for causal emergence does not track ‘causation’ of some kind, but rather that it is unlikely that this kind of causal emergence is going to satisfy anyone looking for something stronger than mere explanatory emergence. Either interventionism (and thus EI) tracks something like causal explanation rather than causal power, which will not satisfy many traditional emergentists, or else it is likely to suffer from something like the fat-handedness objections outlined above. It is possible that the latter could be blocked (see e.g., Kästner & Anderson, 2018 for recent discussion), but this is very much an ongoing debate, and one that it seems like Hoel in particular will struggle to address, as he explicitly acknowledges that changes at the macro-level will supervene on changes at the micro-level. There are two ways he could respond here: one would be to acknowledge that all causation is at the micro-level, and accept that his account of causal emergence is purely epistemic in character; the other would be to argue that the supervenience of the macro on the micro does not block ‘true’ causal emergence, even given concerns about fat-handedness. Even then, we could ask whether the macro-level causes measured by EI should really be described as ‘emergent’.

4 | IS IT REALLY EMERGENT?

In the literature on emergence, it is typical to distinguish between ontological (or strong) emergence and epistemic (or weak) emergence (see e.g., Bedau, 1997, 2002; Chalmers, 2006; Silberstein & McGeever, 1999; Wilson, 2015). This distinction is not entirely uncontroversial, and in particular there might be reason to further distinguish between strong and weak varieties of ontological emergence, etc., but for my purposes the ontological/epistemic distinction will suffice. Ontological emergence, in the sense that matters here, concerns the emergence of genuinely novel properties at some non-fundamental level, while epistemic emergence concerns the emergence of greater explanatory or predictive power at some higher level of description.

Hoel's emergence might initially seem to be clearly epistemic, as it simply gives us a more informative description of coarse-grained processes that could also be described in fine-grained terms at the lower level, without introducing any genuinely novel properties or powers. Hoel himself seems to recognise this, writing that “a macroscale description of a system (a map) can be more informative than a fully detailed microscale description of the system (the territory)” (2017: p. 187, emphasis added). Understood in this way, Hoel's claim is just that the higher level description can be more informative (in the sense measured by EI), not that it introduces anything that is not already present at the lower level.

However it seems like Hoel is also committed to the emergence of novel causal powers at the macro-level, writing that causal emergence “is when the macro beats the micro in terms of efficacy, informativeness, or power of its causal relationships” (2017, emphasis added). It's possible that Hoel just has in mind greater explanatory power here, and thus emergence in an epistemic sense.⁴ If this is the case then the account is less exciting than some have taken it to be, as it would not provide a demonstration of the ontological emergence of novel causes. To say that novel causal powers can emerge at a higher level of organisation seems to be saying something metaphysically robust, and to entail the addition of something genuinely new to our ontology, not just the addition of explanatory or predictive power. Ontological emergence could guarantee the autonomy of mental states and processes from their underlying physical substrates, and thus offer a foundation for the emergence of genuinely novel mental causes of the kind that Kim’s causal exclusion argument is meant to rule out.
Unfortunately it does not seem likely that Hoel’s approach to measuring causal power can achieve ontological emergence of this kind. Recall that he is committed to something like an interventionist approach to causation, where our measure of causal power depends on an intervention into the causal structure of a system. According to Hoel’s approach, if a macroscale intervention is more informative (or predictive) than a microscale one, then there is greater causal power at the macro-level of the system than at the micro-level, demonstrating causal emergence. But any intervention at the macroscale is also going to involve an intervention at the microscale that it supervenes on, and so any higher-level ‘causation’ is also going to be accompanied by a change in lower-level causal dynamics. So even if it is true that we can more reliably predict the effects of macro-level interventions, this does not mean that there is necessarily anything ‘more’ going on at the macro-level (causally speaking), rather just that our macro-level description of a system partitions states in such a way that it is easier to predict the effects of an intervention. Even Woodward’s (2015b) defence of the interventionist response to causal exclusion arguments does not claim that a mental state might possess causal powers that its physical supervenience base lacks, but rather just that both the mental and the physical states might be causes of some subsequent mental state. While this kind of response might be able to preserve mental causation alongside physical causation, it has nothing to say about the emergence of novel mental causes. At best, it seems that Hoel has established the epistemic emergence of higher-level causes, in the sense that we might be able to generate better causal explanations by targeting our interventions at the macroscale of the kind of system he describes.

Hoel’s measure of emergence could still be interesting even if it were merely epistemic. Bedau (2002) describes how weakly emergent causal properties might exhibit “explanatory autonomy and irreducibility”, while nonetheless being “dependent on and reducible to their underlying phenomena” (ibid.: pp. 12–13). The idea is that there might be nothing more to some systems than their underlying micro-states, at least ontologically speaking, and yet we might nonetheless only be able to explain the behaviour of such systems in terms of their higher level macro-states, due to our own epistemic or computational constraints. It could make sense to describe these higher level explanations in causal terms, even if there is strictly speaking no ‘additional’ causal power at the macro-level. Weakly emergent causation of this kind might not even require that we are totally unable to give an explanation at the micro-level, but only that the macro-level explanation is more informative (cf. Taylor, 2015). If Hoel has provided only a formal approach to measuring the sense in which a higher level description of a system can be more informative, he will still have made a valuable contribution (cf. Beckers & Halpern, 2019 for another approach to this kind of question). By measuring the strength of an intervention in terms of effective information, he gives a helpful analysis of the formal circumstances under which a coarse-grained description of a system might be more (epistemically) beneficial than a fine-grained one, and consequently an analysis of the kinds of systems where an epistemically non-reductive ‘special science’ approach is most appropriate. In future work this kind of analysis could be applied more widely to questions to do with the (dis)unity of the sciences and the relationship between different scales of scientific explanation, as discussed by, for example, Fodor (1974) and Ross et al. (2007), even if it cannot provide a robust account of causal emergence in the traditional sense.

5 | CONCLUSION

I have presented Hoel’s proposal for a new model of causal emergence, based on a formal measure of ‘effective information’, and argued that it is causal only in the somewhat deflationary
interventionist sense, and thus emergent only in a weak or epistemic sense. Anyone hoping for a demonstration of the strong emergence of ontologically novel causal powers should therefore not be too excited about this new information-theoretic proposal. Nonetheless, Hoel has provided a valuable analysis of the circumstances under which a macroscale description is more informative about the underlying causal dynamics of a system, and thus a potential measure of causal emergence in the weak, epistemic sense. My aim in this paper was not to intervene in metaphysical debates about the nature of causation, but rather just to clarify what kind of causal emergence is demonstrated by Hoel’s proposal, in order that those who might be looking for something stronger are not misled or disappointed by what is on offer here.

ACKNOWLEDGMENTS
I would like to thank Matteo Grasso, Cecily Whitely, David Silverman, Jonny Lee, Krzysztof Dolega, Natesh Ganesh, Tyler Millhouse, and audience members at the Second Joint MCMP-Hannover Workshop “Philosophy of Science” and the 2019 Joint Session for their helpful comments, suggestions, and encouragement at various stages of this project. I am grateful to the editors and anonymous reviewers of this journal for pushing me to clarify several points, while also highlighting why this novel approach to causal emergence should be of interest to a philosophical audience.

ORCID
Joe Dewhurst https://orcid.org/0000-0002-3497-2508

ENDNOTES
1 For the latest version of this theory (‘IIT 3.0’), see Oizumi et al. (2014).
2 Hoel et al. (2016) present another measure of causal emergence, $\phi^{\text{Max}}$, which draws more on the integrated information theory mentioned above. It imposes stronger requirements than EI, but continues to rely on coarse-graining, which means that much of what I have to say here will still apply. In future work I would like to explore the differences between EI and $\phi^{\text{Max}}$ in greater detail.
3 List and Pivato (2015, see also List, 2014, 2019) have argued for something like the opposite of Hoel’s causal emergence: the emergence of macroscale indeterminacy from a determinate microscale system. Similarly to Hoel, though, they advocate an ontic interpretation of this kind of emergence, despite its apparently epistemic appearance. Gebharter (2020) has recently responded to their argument, using a simple probabilistic model to demonstrate why it does not achieve what it sets out to do.
4 I thank an anonymous reviewer for suggesting this interpretation.

REFERENCES
Albantakis, L., Marshall, W., Hoel, E., & Tononi, G. (2019). What caused what? A quantitative account of actual causation using dynamical causal networks. *Entropy, 21*(5), 459.
Andersen, H. K. (2017). Patterns, information, and causation. *The Journal of Philosophy, 114*(11), 592–622.
Baumgartner, M., & Gebharter, A. (2016). Constitutive relevance, mutual manipulability, and fat-handedness. *British Journal for the Philosophy of Science, 67*, 731–756.
Beckers, S., & Halpern, J. Y. (2019). Abstracting causal models. *Proceedings of the 33rd AAAI Conference on Artificial Intelligence.*
Bedau, M. A. (1997). Weak emergence. *Philosophical Perspectives, 11*, 375–399.
Bedau, M. A. (2002). Downward causation and the autonomy of weak emergence. *Principia Revista Inernacional de Epistemologica, 6*, 5–50.
Braddon-Mitchell, D. (2017). The glue of the universe. In H. Beebee, C. Hitchcock, & H. Price (Eds.), *Making a difference: Essays on the philosophy of causation.* OUP.
Chalmers, D. (2006). Strong and weak emergence. In P. Clayton & P. Davies (Eds.), *The re-emergence of Emergence*. OUP.

Crutchfield, J. (1994a). Is anything ever new? Considering emergence. In G. Cowan, D. Pines, & D. Melzner (Eds.), *Integrative themes*. Addison-Wesley.

Crutchfield, J. (1994b). The calculi of emergence: Computation, dynamics, and induction. *Physica D*, 75, 11–54.

Flack, J. C. (2017). Coarse-graining as a downward causation mechanism. *Philosophical Transactions of the Royal Society A*, 375, 20160338.

Fodor, J. A. (1974). Special sciences (or: The disunity of science as a working hypothesis). *Synthese*, 28(2), 97–115.

Gebharter, A. (2020). Free will as a higher-level phenomenon? *Thought: A Journal of Philosophy, 9*(3), 177–187.

Gillett, C. (2019). Emergence, downward causation and its alternatives: Critically surveying a foundational issue. In S. Gibb, R. F. Hendry, & T. Lancaster (Eds.), *The Routledge handbook of emergence*. Routledge.

Griffiths, P., Pocheville, A., Calcott, B., Stotz, K., Kim, H., & Knight, R. (2015). Measuring causal specificity. *Philosophy of Science*, 82, 529–555.

Hoel, E. (2017). When the map is better than the territory. *Entropy*, 19, 188. https://doi.org/10.3390/e19050188

Hoel, E., Albantakis, L., Marshall, W., & Tononi, G. (2016). Can the macro beat the micro? Integrated information across spatiotemporal scales. *Neuroscience of Consciousness*, 2016(1), niw012. https://doi.org/10.1093/nc/niw012

Hoel, E., Albantakis, L., & Tononi, G. (2013). Quantifying causal emergence shows that macro can beat micro. *Proceedings of the National Academy of Science*, 110, 19790–19795.

Kästner, L., & Anderson, L. M. (2018). Intervening into mechanisms: Prospects and challenges. *Philosophy Compass*, 13, e12546.

Kim, J. (2000). *Mind in a physical world*. MIT Press.

Kim, J. (2005). *Physicalism or something near enough*. Princeton University Press.

List, C. (2019). Free will, determinism, and the possibility of doing otherwise. *Noûs*, 48(1), 156–178.

List, C. (2019). *Why free will is real*. HUP.

List, C., & Pivato, M. (2015). Emergent chance. *Philosophical Review*, 124(1), 119–152.

Oizumi, M., Albantakis, L., & Tononi, G. (2014). From the phenomenology to the mechanisms of consciousness: Integrated information theory 3.0. *PLoS Computational Biology*, 10(5), e1003588.

Owen, M. (2019). Exploring common ground between integrated information theory and aristotelian metaphysics. *Journal of Consciousness Studies*, 26(1–2), 163–187.

Pearl, J. (1988). *Probabilistic reasoning in intelligent systems*. Elsevier.

Pearl, J. (2000). *Causality: Models, reasoning, and inference*. CUP.

Reutlinger, A. (2012). Getting rid of interventions. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 18(4), 787–795.

Robb, D. (2019). Emergent mental causation. In S. Gibb, R. F. Hendry, & T. Lancaster (Eds.), *The Routledge handbook of emergence*. Routledge.

Rosas, F. E., Mediano, P. A. M., Jensen, H. J., Seth, A. K., Barent, A. B., Carhart-Harris, R. L., & Bor, D. (2020). Reconciling emergences: An information-theoretic approach to identify causal emergence in multivariate data. *PLoS Computational Biology*, 16(12), e1008289.

Ross, D., Ladyman, J., & Collier, J. (2007). Rainforest realism and the unity of science. In J. Ladyman & D. Ross (Eds.), *Every thing must go*. OUP.

Seth, A. (2010). Measuring autonomy and emergence via Granger causality. *Artificial Life*, 16(2), 179–196.

Shapiro, L., & Sober, E. (2007). Epiphenomenalism – The do’s and don’t’s. In G. Wolters & P. Machamer (Eds.), *Thinking about causes*. University of Pittsburgh Press.

Silberstein, M., & McGeever, J. (1999). The search for ontological emergence. *The Philosophical Quarterly, 49*(195), 201–214.

Spirtes, P., Glymour, C., & Scheines, R. (1993). *Causation, prediction, and search*. Springer.

Strevens, M. (2007). Review of Woodward, *Making things happen*. *Philosophy and Phenomenological Research, 74*(1), 233–249.

Strevens, M. (2008). Comments on Woodward, *Making things happen*. *Philosophy and Phenomenological Research, 77*(1), 233–249.

Taylor, E. (2015). An explication of emergence. *Philosophical Studies, 172*, 653–669.
Tononi, G. (2008). Consciousness as integrated information: A provisional manifesto. *Biological Bulletin, 215*, 216–242.

Tononi, G., & Sporns, O. (2003). Measuring information integration. *BMC Neuroscience, 4*, 31.

Wilson, J. (2015). Metaphysical emergence: Weak and strong. In T. Bigaj & C. Wuthrich (Eds.), *Metaphysics in contemporary physics. Poznan studies in the philosophy of the sciences and the humanities*. BRILL/RODOPI.

Woodward, J. (2003). *Making things happen*. OUP.

Woodward, J. (2008). Response to Strevens. *Philosophy and Phenomenological Research, 77*(1), 193–212.

Woodward, J. (2015a). Methodology, ontology, and interventionism. *Synthese, 192*, 3577–3599.

Woodward, J. (2015b). Interventionism and causal exclusion. *Philosophy and Phenomenological Research, 91*(2), 303–347.

**How to cite this article:** Dewhurst J. Causal emergence from effective information: Neither causal nor emergent? *Thought: A Journal of Philosophy*. 2021;1–11. [https://doi.org/10.1002/tht3.489](https://doi.org/10.1002/tht3.489)