TOPICS
TOPICS IN COGNITIVE SCIENCE

Topics in Cognitive Science (2019) 1–16

ISSN:1756-8765 online
DOI: 10.1111/tops.12445

This article is part of the topic "Levels of Explanation in Cognitive Science: From Cultures to Molecules," Matteo Colombo and Markus Knauff (Topic Editors). For a full listing of topic papers, see http://onlinelibrary.wiley.com/journal/10.1111/(ISSN)1756-8765/earlyview

Experiential Explanation

Sara Aronowitz,[a] Tania Lombrozo[b]

[a]*University Center for Human Values & Program in Cognitive Science, Princeton University*
[b]*Department of Psychology, Princeton University*

Received 15 February 2019; received in revised form 28 June 2019; accepted 28 June 2019

Abstract

People often answer why-questions with what we call experiential explanations: narratives or stories with temporal structure and concrete details. In contrast, on most theories of the epistemic function of explanation, explanations should be abstractive: structured by general relationships and lacking extraneous details. We suggest that abstractive and experiential explanations differ not only in level of abstraction, but also in structure, and that each form of explanation contributes to the epistemic goals of individual learners and of science. In particular, experiential explanations support mental simulation and survive transitions across background theories; as a result, they support learning and help us translate between competing frameworks. Experiential explanations play an irreducible role in human cognition—and perhaps in science.

Keywords: Explanation; Abstraction; Narrative; Storytelling; Learning; Mental simulation; Theory change

1. Introduction

Why did you move to New Jersey? Compare two answers to this question. "Because the recession increased unemployment in Michigan, but the industry was expanding on the East Coast." "Because one day, I ran into my friend at a bar, and she told me she had been laid off; after that I spent a lot of time worrying and started looking for other jobs just in case and eventually came across a great opportunity in New Jersey." These

Correspondence should be sent to Sara Aronowitz, Department of Psychology, Princeton University, Princeton, NJ 08540. E-mail: skaron@umich.edu

responses both explain why a particular event[1] occurred, but they have fundamentally different structures. The first answer appeals to general facts, which are linked to the particular instance being explained (the explanandum). The second answer appeals to a particular series of events in a temporal order, and involves much more detail.

We will refer to explanations of the first kind as *abstractive explanations*: They feature the most general structure that still differentiates the explanandum from relevant alternatives and satisfies other explanatory criteria, such as simplicity and completeness. This means that the structure of abstractive explanations will vary depending on what is being explained: In the New Jersey case, the structure centers around economic trends, but with a different explanandum, an abstractive explanation could feature cultural shifts, urban development, and so on. Abstractive explanations (suitably refined) seem to be what philosophers of science such as Kitcher (1985) or Lange (2016) have in mind when they discuss explanations, and it is often what psychologists have in mind as well (e.g., Keil, 2006; Lombrozo, 2012; Wellman, 2011).

We will call answers of the second kind *experiential explanations*. These explanations have in common an experiential structure; that is, they are structured around dimensions that apply to our ordinary experience, such as time, space, and sensory modality. As we use the term, experiential explanations need not preserve *all* of these dimensions, but they must use at least one. Experiential explanations can be excessively detailed, but in general they do not seem worse for the inclusion of details that are not necessary to structure the event (see Vossen, Caselli, & Cybulska, 2018 for a discussion of detail in stories). For example, the inclusion of location (a bar) does not seem to diminish our introductory experiential explanation, whereas mentioning that one learned about the recession in a bar does seem to weaken the abstractive explanation. Experiential explanations include what psychologists and philosophers have called narratives (Bruner, 1991; Currie, 2007), case studies (Kagan, 2001), and stories (Gelman & Basbøll, 2014).

Rather than being two mutually exclusive categories, real-world explanations often involve a mixture of abstractive and experiential features. For example, our introductory example could begin with the friend at the bar, but appeal to economic trends as well. Even if real-world cases are often mixed in this way, we can distinguish between clear cases of each explanation type, and our analysis will proceed accordingly. Specifically, we will take (at least partial) subsumption under an abstract structure to be constitutive of abstractive explanation and focus our attention on two features of experiential explanations that differentiate them from their abstractive counterparts: their experiential structure and their permissiveness when it comes to concrete detail (see Table 1).

While abstractive explanations are widely acknowledged to be critical to both scientific progress and learning (Lombrozo, 2011), the status of experiential explanations is less clear. In the context of science, experiential narratives or stories are often derided as "anec-data," and they are taken to diverge from scientific norms (Sloman,; Bruner, 1991; Rosenberg, 2018). Nonetheless, we contend that experiential explanations are indeed *explanations*. Not only are they answers to why questions, they are also designed to satisfy the explanatory demand behind a why-question in promoting understanding of the explanandum.[2] Experiential explanations thus differ from what we might call

Table 1
Properties of abstractive and experiential explanations

	Abstractive Explanations	Experiential Explanations
Structuring dimensions	Determined contextually by principles/ regularities (e.g., mass, location, GDP, population size . . .)	One or more of a set of experiential dimensions (time, space, sensory modality, affect . . .)
Level of detail	As minimal as possible while satisfying other explanatory virtues	Permissive; constrained by narrative or story-telling conventions
Schema	1. *General principle(s)/regularities* 2. Relevant features of the case at hand structured according to the general principles/regularities 3. *Proposition linking 1&2 to the explanandum*	1. *Narrative frame, which specifies an episode and perspective* 2. Relevant features (plus potentially tangential features) of the case at hand structured accorded to the narrative frame 3. *Proposition linking 1&2 to the explanandum*
Example	1. Human migration is often spurred by economic factors 2. New Jersey had lower unemployment at the time in my industry 3. I moved because of increased economic opportunity in New Jersey	1. My perspective in the weeks before the move 2. Meeting with a friend at a bar, hearing about her complaint, returning to work, deciding 3. I moved because of these experiences

Note. In the schemas, italics denote content that is sometimes only implicit.

"experiential descriptions," which present information about an event without organizing it to answer a why-question.

Our aim in this paper is to mount a defense of the rationality of experiential explanations. We will argue that experiential explanations have a vital role in learning because of their unique structure and level of detail.[3] Experiential explanations are better suited than abstractive explanations to function as inputs to mental simulation and can carry over to a new context when background theory changes, making old theories and evidence intelligible from a new standpoint. We conclude with a brief discussion of how our argument about individual thinkers applies to science and scientific psychology.

2. Advantages of abstractive explanation

Abstractive explanations, and the process of abstraction itself, are widely thought to be the building blocks of scientific knowledge, as well as everyday understanding of the world. On Friedman's (1974) view, for example, a successful explanation reduces the number of independent facts we need to accept; such explanations are abstractive since they are only explanatory insofar as they involve a reduction of particular details to laws, regularities, or other generalizations. Other dominant views of explanation, while departing from Friedman's, share a commitment to abstraction. For example, Strevens's (2004) Kairetic account treats generality as a desideratum, where generality is a matter of the causal model being applicable to as many cases as possible.

Like philosophers, psychologists highlight the value of abstraction in explanation and beyond. In a review of psychological approaches to abstraction, Burgoon, Henderson, and Markman (2013) cite advantages of abstraction in prediction, memory, and decision-making. For example, more abstract/ less detailed representations can help children appreciate relational properties (e.g., Gentner & Ratterman, 1991), and they are a hallmark of expertise in adults (e.g., Hinds, Patterson, & Pfeiffer, 2001; Rabinowitz & Friedman, 2018). Moreover, research on the role of explanation in learning has been used to argue that seeking and generating explanations can itself promote abstraction, which can have positive downstream consequences, such as supporting transfer and subsequent learning (Walker & Lombrozo, 2017; Wilkenfeld & Lombrozo, 2015; Williams & Lombrozo, 2010).

Abstraction, as understood in philosophy and psychology, is a way of getting at general features of our environment and coming to understand not just what is the case, but how various facts interact and depend on one another. This modal understanding could be intrinsically valuable, but also enable accurate generalization, since generalization involves determining which features make a difference in producing a state of affairs, how they do so, and which features are irrelevant.

3. Advantages of experiential explanation

Despite the clear benefits of abstraction, there is evidence that we do not always prefer explanations with abstract structures and minimal extraneous detail. Bechlivanidis et al. (2017), for example, found that their participants favored explanations with more precise, concrete details, though irrelevant details did not have added value (see also Frazier, Gelman, & Wellman, 2016). More generally, there's evidence that when engaged in activities that might broadly be considered explanatory, people offer and accept "narratives" or "stories" that incorporate concrete and temporal information (e.g., Bruner, 1991; Dawes, 1999; Snow & Beals, 2006; see also Schechtman, 2007). This naturally raises the question: Why don't we always engage in abstractive explanation? Or, put another way: There may be social, aesthetic, or other pragmatic reasons for engaging in experiential explanation, but are there *epistemic*[4] reasons to do so? If, following Friedman, we accept that there is a rational, epistemic pressure to favor abstraction, is there a countervailing rational, epistemic pressure to favor explanations with experiential structure and concrete detail?

Perhaps we do not always engage in maximally abstractive explanation because different *levels* of abstraction have their own advantages. Within philosophy, Weatherson (2012) introduces the "Goldilocks Problem"—the idea that a good (abstractive) explanation should balance between too much and too little detail.[5] Even mechanistic theories of explanation, which some have characterized as too lenient toward the inclusion of concrete details, acknowledge a role for abstraction in addition to completeness (Craver & Kaplan, 2018). Within psychology, Burgoon et al. note that "common to each of the subdisciplines of psychology is the notion that processing information *at different levels of abstraction* is functional" (emphasis added; see also Trope & Liberman, 2010). Thus, even within the category of abstractive explanations, there's reason to expect variation in *level* of abstractness.[6]

On our view, what makes an explanation abstractive is not that it is *maximally* abstract, but rather that it is explanatory by virtue of relating the explanandum to a structure that is more abstract than the explanandum. Thus, experiential explanations—as we have defined them—are not just at a different or more concrete level than abstractive explanations. Experiential explanations also have a structure that is in some sense isomorphic to the structure of actual experience.[7] What, then, are the advantages of experiential explanation? In the two sections that follow, we suggest that two key features of experiential explanations—experiential structure and permissive degree of concrete detail—combine to support important epistemic advantages.

3.1. Inputs for simulation

The first advantage of experiential explanations comes from their role as inputs for mental simulation. Mental simulation encompasses a heterogeneous set of activities—such as projecting the consequences of an arm movement, or imagining how you would feel if an event in the news happened to you—that share an isomorphic mapping to the episodes they simulate. Just as an arm movement (or a newsworthy event) unfolds over time and from a point of view, so, too, the corresponding mental simulation has a temporal structure and is experienced from a perspective.

Scholars differ in how they define mental simulation. For example, Spaulding (2016) defines simulation as involving a concrete replication: Simulation seems like experiencing or "bringing to life" because it functions by replicating to some extent the computations involved in actual experience. In the case of theory of mind, a particular (and controversial) view of how this ability is implemented is the Simulation Theory (Goldman, 1992; Gordon, 1986). As we explain further below, our appeal to "simulation" has more modest commitments—a key feature is that simulation generates a fairly specific token episode or set of such episodes which in some sense resemble an episode that could be experienced. There need not be any genuinely experiential *computation* behind this resemblance, such as a reuse of sensory capacities. Our argument, in brief, will be that experiential explanations provide better inputs to mental simulation than do abstractive explanations because they preserve the relevant experiential structure to trigger (accurate) simulations.

As an illustrative example, consider a mental simulation of how an acquaintance felt when she attended a beach wedding in a bathing suit, only to discover the other attendees in formal attire. Putting yourself in her shoes (or her flip flops, as the case may be), you would likely come to believe that she felt embarrassed. This conclusion depends not only on performing the right simulation (that is, using accurate and relevant aspects of the story as "inputs" to your simulation), but also on using inputs of the right *kind*. If the scenario is described much more abstractly (as, say, an episode in which an acquaintance recognized a false belief that led her to publicly violate a social expectation or norm), the inference to *embarrassment* follows much less readily. This idea is not restricted to simulations in the domain of theory of mind; in some mechanical reasoning tasks, for example, people are more likely to generate accurate predictions when they generate visuomotor simulations of the relevant events, as if they were actually performing or observing them (Schwartz & Black, 1999).[8]

These examples suggest a more general and intuitive empirical principle, which we'll call *isomorphic input*: that mental simulations will be triggered more often and yield more reliable conclusions when their inputs resemble the actual experiences they are simulations of. This principle could be supported by a strong commitment to simulation as some form of neural/ cognitive re-use: If mental simulations re-use the same mechanisms employed in first person experience, we might expect experiential inputs to provide a better match to the proper inputs for such mechanisms. But the principle is also supported by a much weaker commitment, namely that first person experience supports many forms of learning (motor, visual, conceptual, etc.), where different kinds of inputs are more or less effective at tapping into each form of learning. For example, being walked through an emotional experience from a particular perspective could tap into associative learning about emotions; performing a visual simulation could tap into mechanical principles implicitly encoded in statistical learning mechanisms. We suggest that experiential inputs are better at tapping into the knowledge implicitly or explicitly encoded in many such mechanisms, and that this is in virtue of their temporal/ sensory/ perspectival structure.

For all we have said so far, experiential *descriptions* should be just as effective at triggering mental simulations as experiential *explanations*. While both may indeed trigger simulations, only experiential explanations are linked to and organized as responses to why-questions. As a result, they will work better as inputs to mental simulations targeted at answering those why-questions than will corresponding experiential descriptions, which would require search and organization to serve an equivalent purpose. For instance, in the wedding case, one considers the situation from the perspective of the acquaintance to understand *why she felt embarrassed at a specific point in time*. One could instead be interested in why she felt enthusiastic earlier in the day, or why the other guests did not mention her outfit, and so on; answering these questions would most naturally employ a different simulation, whether at a different time or from a different perspective.

The role of experiential explanation in mental simulation is supported by the isomorphic input principle, but why expect the principle to hold? The evidence is compelling but indirect. One source of evidence comes from language comprehension. Richardson and Matlock (2007), for example, found that the language used to describe a proposition (for instance, "the road *runs* through the valley" vs. "the road *is in* the valley") influenced the extent to which participants generated eye movements associated with the simulation of motion (in this case *running*), consistent with the experiential properties of the figurative language employed (see also perceptual simulation theory, Zwaan, Madden, Yaxley, & Aveyard, 2004). Experiential language can also trigger simulations involving other modalities: Kurby and Zacks (2013) found correlations between varieties of experiential language and neural activity associated with corresponding visual, auditory, and motor areas. Findings like these suggest that explanations with more concrete/ modally rich language will trigger more simulation.

Another source of evidence comes from research on text comprehension, which suggests that in both reading and generating narratives, people construct "situation models" with characteristic temporal, causal, and intentional structure (Zwaan, Langston, & Graesser, 1995). With respect to temporal order, studies suggest that when evidence concerning some event preserves temporal structure, it is found more persuasive (Pennington & Hastie, 1988). Narratives also tend to employ more concrete language (Costabile, 2016), which is associated with ease of mental imagery—itself a form of simulation. To the extent experiential explanations incorporate similar temporal structure and concrete language, we should expect them to better support the functions typically attributed to such models of evidence and text comprehension—including the ability to draw inferences beyond what is stated explicitly (e.g. Sadoski,1983).

More generally, the *capacity* to represent and generate experiential explanations seems to be in some way connected to the ability to simulate. For instance, there is a well-documented connection between episodic memory and imaginative future thinking, including a shared neural basis for both in the hippocampus (Campbell, Benoit, & Schacter, 2017; Parikh, Ruzic, Stewart, Spreng, & De Brigard, 2018). One way of reading this connection is what Schacter and Addis (2007) call the *Constructive Episodic Simulation Hypothesis*, on which a single cognitive system is responsible for constructing past scenes, future predictions, and imagined alternative possibilities. On this hypothesis, details of the past structured in an experiential, episodic form are utilized in imaginative simulation more generally. As such, there may be a connection both between experiential information and simulation, and between the capacity to construct experiential episodes and the capacity to project hypothetical possibilities.

So far we have suggested a mapping between the contents of experiential explanations, on the one hand, and effective inputs to trigger or carry out mental simulations, on the other. But it is a further step to suggest that this relationship to mental simulation has epistemic advantages. After all, it could be that more concrete (and emotionally laden) stimuli result in worse decisions (e.g., Keysar, Hayakawa, & An, 2012), or that theory of mind simulations *seem* reliable while actually being biased or shallow (Rosenberg, 2018). We certainly won't make the case that mental simulation is always accurate or beneficial. Nonetheless, we will suggest that in some cases of interest for both everyday cognition and science, experiential explanations have benefits that are not subsumed by their abstractive counterparts.

First, there is clear evidence of mnemonic benefits for more concrete (as opposed to abstract) material. Sadoski, Goetz, and Rodriguez (2000), for example, found that participants were more likely to recall the gist of prior text when it was more concrete versus abstract (defined in terms of ease of forming a mental image). Moreover, concreteness is associated with both self-reported comprehensibility (Sadoski, Goetz, & Fritz, 1993) and performance on objective measures of comprehension (Sadoski, 1985; Wharton, 1980). These advantages for memory and comprehension are clearly valuable in themselves, but they might also improve our ability to evaluate and revise such explanations over time.

Second, a large body of research confirms that more concrete representations (which we contend better support simulations) yield different patterns of judgments. For example, manipulations of concreteness have been shown to shift moral judgments (Nichols & Knobe, 2007) and social inferences (Costabile, 2016). In these domains, it is unclear whether the concrete representation results in more accurate or beneficial judgments, but in other domains with objectively correct answers (such as mechanical reasoning), more concrete/experiential representations can yield the more accurate response (Schwartz & Black, 1999). Moreover, there is evidence that in some cases, mentally simulated experience can serve as a substitute for actual experience—for example, mental practice of a piano performance had a comparable effect as physical practice in expert pianists (Bernardi, Schories, Jabusch, Colombo, & Altenmuller, 2013; see Kappes & Morewedge, 2016 for an overview).

In sum, experiential explanations sometimes trigger mental simulation and, in so doing, support inferences that would otherwise be less memorable, less comprehensible, and/or less accessible. These properties derive in large part from the structure and detail of experiential explanations, and how these elements interact with language comprehension, the peculiarities of our simulatory capacities, and the organization of human memory.

3.2. Repurposing

Consider the following use of experiential explanation. Patient H.M. was introduced into memory research by Scoville and Milner (1957), at a time when researchers were pessimistic about localizing memory to a region of the brain.[9] They originally saw H.M.'s case as illustrating the importance of the hippocampus in memory, writing: "it is concluded that the anterior hippocampus and hippocampal gyrus, either separately or together, are critically concerned in the retention of current experience" (21). Subsequently, memory research focused increasingly on the role of the hippocampus, with some researchers going as far as suggesting that episodic memory was localized to the hippocampus. However, later study of the same case revealed that H.M. had damage outside of the hippocampal area, and that many of the deficits observed in H.M. (but not in other patients with similar lesions) could be linked to this damage. Across time, H.M.'s case featured in several importantly different explanations of why his brain damage caused a particular pattern of memory deficits, and more generally, of why MTL brain damage affects memory.

Scoville and Milner in fact provided two explanations of the memory phenomena observed in H.M. The first was an abstractive explanation that linked hippocampal damage to episodic deficits, structured around generalizations about which brain regions support which kinds of functioning. The second was an experiential explanation about observations of H.M.'s behavior and pattern of damage, structured around a series of temporally-indexed and particular findings; for example, they contrast the way H.M.'s deficits were obvious to his family after a move with the way he "seems like a relatively normal individual" to a casual observer (14). In this case, the abstractive explanation is of merely historical interest from our current vantage point. On the other hand, the experiential explanation still explains something about the neural basis of memory as well as the connection between H.M.'s episodic deficits and his medial temporal lobe damage. This section explores whether the *survival* of experiential explanations across episodes of scientific change—be it through direct repurposing or partial co-option—is an epistemic advantage. We will consider two versions of the view that repurposing serves as a kind of evidence, but ultimately argue for a less directly evidential role for repurposing: ameliorating the harm of incommensurable transitions in background theory.

An initial possibility is that repurposing is a kind of falsification of the original theory: Co-opted experiential explanations can help us see where our theories went wrong. Gelman and Basbøll (2014) offer a defense of storytelling in science that develops a related idea. They describe a case with the same structure as the H.M. example, where a story presented in support of one view of human nature can be reused to criticize the very same view. While in no sense taking the place of statistical evidence, this story still occupies an evidential role in exposing the weaknesses of a theory. To be successfully repurposed, Gelman and Basbøll argue, a story must be immutable (i.e., fixed in some determinate facts, rather than flexible according to the whims of the storyteller) and anomalous (i.e., contain some information that is not well explained by existing theories). Both of these features hold in the case of H.M.: His story was fixed by actual facts that allowed it to be reinterpreted as more was discovered about the extent of his injuries, and it also initially contradicted the existing theories in the 1950s and earlier, within which memory was taken to rely on widely distributed processing.

On a strong version of the view that experiential explanations play an evidential role, experiential explanations allow us to go beyond our current thinking, which is constrained by our current theories. As compared to abstractive explanations, which are heavily theory-laden, experiential explanations may be *less* beholden to existing frameworks. However, thinking carefully about how experiential explaining works raises issues for this idea, as experiential explanations themselves involve a great deal of theory-laden selection.

An initial form of selection occurs in choosing a particular narrative frame in response to a why-question: We have failed to provide an explanation at all if we do not find an episode and perspective that is relevant to our theory and its competitors, and so this selection will be heavily determined by what we see as the space of relevant possible theories. A second form of selection occurs over facts within the narrative frame. These forms of selection are necessary: They are what distinguishes experiential explanation from mere description. The second form of selection may be *less* theory-laden than the first, since details are often chosen based on narrative conventions and the demands of an experiential structure rather than anything about the specific theory at hand. Nonetheless, these selection processes indicate that experiential explanations are still theory-laden, albeit in a different way than abstractive explanations.

Accepting these relationships between theory and experiential explanation still allows for a weak version of the Gellman and Basbøll view. An optimistic reading is that if the old experiential explanation can be successfully reworked under the new theory, then the new theory must have excellent explanatory resources, since it can account for a case that was originally selected on the basis of its support for the old theory. However, the old theory and the new theory may not be the only options that are or should be on the table, and so it may be that the repurposing is easier because of shared ground between the old and new theory. Here, repurposing allows traces of the old theory to carry forward, and privileges commonalities between the new and old theory.

Notice that the description of how H.M.'s case was repurposed over time is itself an experiential explanation of how the theory change took place. This explanation helps illuminate the transition from an older, cruder form of localization to a more nuanced version, and it forms a blueprint that might be applied to other cases—allowing us, for instance, to look back on other lesion studies and attempt to repurpose them via an analogous explanatory structure. This brings us to an advantage of even simple reuse. Even if repurposing cannot solve the problem of theory-laden evidence, all kinds of reuse are a kind of translation that helps us solve a different epistemic problem: incommensurability.

The problem of incommensurability has been raised in the scientific context by Kuhn (1962/2012) and developed in a more personal context by Paul (2014). For our purposes, it's enough to acknowledge incarnations of modest incommensurability that raise practical concern in changes in explanatory structures. Kuhn describes the practitioner of normal science after a paradigm shift, claiming that "the world of his research will seem, here and there, incommensurable with the one he had inhabited before. That is another reason why schools guided by different paradigms are always slightly at cross-purposes" (112). That is, at least when we shift between sufficiently different theories, we face a difficulty understanding our past views, in particular a difficulty in comparing our current views to our past views as subsumed under a scientific aim or question. This difficulty is both intrinsically bad, in the way that lack of understanding is always epistemically bad, and also instrumentally bad, since it prevents us from effectively utilizing past evidence.

We suggest that the reuse of experiential explanations can ameliorate the problem of incommensurability.[10] When an experiential explanation is reused, it is by definition intelligible under both theories and coordinates both theories with respect to a single why-question. And as we have argued, experiential explanations are uniquely reusable. With these considerations in mind, we can observe a few contexts in which experiential explanations should be especially useful: in facilitating communication and cooperation across individuals[11] with different background theories (or the same individual over time), and in explaining a case for which we still lack an adequate abstractive explanation that answers a crucial why-question. In each case, an experiential form of explanation will package information in a way that can be useful when translating across disagreeing frameworks, or taken up before an adequate framework is found. Repurposing, on top of simple reuse, makes the most of this translation—in the first kinds of cases, by allowing the disagreeing parties to mediate disputes by returning to the original case and finding new details. In the second kind of case, repurposing allows relevant abstractive dimensions to emerge later and yet still guide some of the selection of information. In our example, Scoville and Milner's original observations were subsequently reanalyzed under new theories, such as functional connectivity or the involvement of the neocortex in remote memory.

In summary, repurposing and reuse do not let us get away from background assumptions into an objective standpoint, but they do allow us to move more easily between competing standpoints and to organize information without the help of a (fully developed) theory.

4. Conclusion

The structure of experiential explanations helps us package information relevant to answering a particular why-question in order to learn better, both in isolation and in combination with abstractive explanations. We have argued that experiential explanations are aptly structured to serve as effective inputs to mental simulation, and we presented some preliminary evidence that these explanations trigger simulatory responses. A second epistemic advantage of experiential explanations is that they can be repurposed under a new theory, offering a bridge across theories that makes them mutually comprehensible, allowing information to be organized, stored, and communicated—even absent a full background theory.

These advantages of experiential explanations follow in large part from the very features that set them apart from abstractive explanations: their experiential structure and their particularity (including the inclusion of concrete detail). It is these features that support mental simulation. Moreover, in repurposing, which details of an explanation might be relevant under the new theory cannot be known in advance or "from the outside"; by not trimming away all particular details, we leave handholds for later repurposing by others or our future selves. Explanations structured around experiential dimensions make weaker commitments concerning which abstractive relationships are explanatory; as such, they remain consistent with more possibilities, and with states of partial ignorance, while still linking a set of facts to a particular why-question and allowing compression and simplification of the episode at hand.

The most flat-footed critique of experiential explanations argues that they are merely "anec-data," or single, improperly sampled cases *acting as data*. Relatedly, Kinzel (2015) lays out a debate on whether case studies can play an evidential role (i.e., act as evidence). However, the preceding discussion reveals that experiential explanations are not substitutes for evidence, but instead help us uptake, communicate, and transform existing evidence. They also go beyond description—they are genuine explanations in answering why-questions to yield explanatory understanding. While previous work has documented several forms of explanatory pluralism (e.g., Colombo, 2017; Lombrozo, 2011, 2012), both descriptively and normatively, such distinctions usually fall within the category of abstractive explanations. The contribution of this paper is in articulating an alternative to abstractive explanation with a unique cognitive and epistemic profile.

Simulation and repurposing surely do not exhaust the role of experiential explanations in learning. Other possibilities to explore include the diversity of explanatory structures aiding in theory search and analogical reasoning, as well as a connection between experiential structure and the move from explanation to intervention.

A final question to consider is whether experiential explanations play any special role within the psychological or social sciences (vs. the natural sciences). One reason to align experiential explanations with the social sciences comes from a long tradition differentiating first-personal or empathic understanding from third-personal or scientific understanding (e.g., Dilthey, 2002; Grimm, 2016; Taylor, 1985). To the extent a discipline such as psychology is inherently perspectival or particular, experiential explanations could play a special role. We are wary of this line of thought; experiential explanations need not involve mental content,[12] and contemporary psychology overwhelmingly aims to achieve *scientific* understanding. On the other hand, as a relatively immature science, and one to whose subject matter we have frequent and direct experiential access, psychology might be well-poised to benefit from mental simulation, and from repurposing as theories evolve.

Acknowledgments

We thank Matteo Colombo, Naftali Weinberger, attendees of the Yale Workshop on Mutual Understanding, and two anonymous reviewers for helpful comments.

Notes

1. While this is a case of explaining a particular event, these two explanatory types can also apply to explanations of general events—for instance, one might explain the acceleration of gravity by reference to other principles of physics (abstractive) or through a narrative about how an object would fall (experiential). Thus, the experiential/abstractive distinction is orthogonal to the generic causation/actual causation distinction. In what follows, we focus on actual explananda for the sake of consistency.
2. Many other ways of thinking of the explanatory demand will be consistent with the argument in this paper.
3. We argue for two roles for experiential explanation that invoke both the structure and level of detail. However, it's beyond the scope of this paper to fully consider the possible function of partially experiential explanations that only have one of these two critical features. Thanks to Naftali Weinberger for suggesting this point.
4. By epistemic reasons, we mean roughly reasons relating to acquiring knowledge or understanding.

5. Some solutions to this problem, such as Blanchard's (2019), will entail that the explanation being as abstract as possible is not a reflection of two explanatory values, abstraction and some countervailing value, but instead a single value. In Blanchard's case, this value is crafting explanations that help us understand all the ways we might intervene.
6. For instance, an ecological explanation of an owl diet might contain less detail about gastrointestinal biology than a physiological explanation (and be more abstract in this sense), but both explain by relating problem-specific structures to more abstract structures, be they relationships between predator and prey or metabolic principles.
7. Plausibly, temporal structure is the most central to experiential explanation; a narrative that describes how the speaker felt at various different times in a sequence without any sensory-modal or spatial detail ("I was first angry, then sad, but then I realized it was a good thing") feels like a perfectly good experiential explanation, even if it fails to convey rich perceptual impressions.
8. Defending the epistemic role of mental simulation is outside the scope of this paper, so we will assume that there is such a role and leave it open whether mental simulation is only a route to learning because of human cognitive particularities, or has a more general epistemic significance. See Lombrozo (in press) for an articulation of one such role.
9. Our depiction of this case draws from Squire (2009).
10. One way to think of this result is as a case of translation between theories that can be accomplished in a restricted context, even given incommensurability (Kitcher, 1982).
11. Along these lines, Mahr and Csipra (2017) argue that episodic memory primarily serves a social, communicative function; on our view, the social context is one among several structurally similar cases where incommensurability may arise, and the problem of incommensurability is broader than the problem of justification that they raise.
12. Consider, for example, an experiential explanation of a geological or evolutionary event, involving a temporal sequence of events from the perspective of a landmass or population.

References

Bechlivanidis, C., Lagnado, D. A., Zemla, J. C., & Sloman, S. (2017). Concreteness and abstraction in everyday explanation. *Psychonomic Bulletin & Review*, *24*(5), 1451–1464. https://doi.org/10.3758/s13423-017-1299-3

Bernardi, N. F., Schories, A., Jabusch, H. C., Colombo, B., & Altenmüller, E. (2013). Mental practice in music memorization: an ecological-empirical study. *Music Perception: An Interdisciplinary Journal*, *30* (3), 275–290.

Blanchard, T. (2019). Explanatory abstraction and the goldilocks problem: Interventionism gets things just right. *The British Journal for the Philosophy of Science*. https://doi.org/10.1093/bjps/axy030

Bruner, J. (1991). The narrative construction of reality. *Critical Inquiry*, *18*(1), 1–21.
Burgoon, E. M., Henderson, M. D., & Markman, A. B. (2013). There are many ways to see the forest for the trees: A tour guide for abstraction. *Perspectives on Psychological Science*, *8*(5), 501–520. https://doi.org/10.1177/1745691613497964
Campbell, K. L., Benoit, R. G., & Schacter, D. L. (2017). Priming, not inhibition, of related concepts during future imagining. *Memory*, *25*(9), 1235–1245. https://doi.org/10.1080/09658211.2017.1283420
Colombo, M. (2017). Experimental philosophy of explanation rising: The case for a plurality of concepts of explanation. *Cognitive Science*, *41*(2), 503–517. https://doi.org/10.1111/cogs.12340
Costabile, K. A. (2016). Narrative construction, social perceptions, and the situation model. *Personality and Social Psychology Bulletin*, *42*(5), 589–602. https://doi.org/10.1177/0146167216636627
Craver, C. F., & Kaplan, D. M., (2018). Are more details better? On the norms of completeness for mechanistic explanations. *The British Journal for the Philosophy of Science*.axy015. https://doi.org/10.1093/bjps/axy015
Currie, G. (2007). Both sides of the story: Explaining events in a narrative. *Philosophical Studies*, *135*(1), 49–63. https://doi.org/10.1007/s11098-007-9092-1
Dawes, R. M. (1999). A message from psychologists to economists: Mere predictability doesn't matter like it should (without a good story appended to it). *Journal of Economic Behavior & Organization*, *39*(1), 29–40.
Dilthey, W. (2002). *The formation of the historical world in the human sciences*. Princeton, NJ: Princeton University Press.
Frazier, B. N., Gelman, S. A., & Wellman, H. M. (2016). Young children prefer and remember satisfying explanations. *Journal of Cognition and Development: Official Journal of the Cognitive Development Society*, *17*(5), 718–736. https://doi.org/10.1080/15248372.2015.1098649
Friedman, M. (1974). Explanation and scientific understanding. *The Journal of Philosophy*, *71*(1), 5–19. https://doi.org/10.2307/2024924
Gelman, A., & Basbøll, T. (2014). When do stories work? Evidence and illustration in the social sciences. *Sociological Methods & Research*, *43*(4), 547–570. https://doi.org/10.1177/0049124114526377
Gentner, D., & Rattermann, M. J. (1991). *Language and the career of similarity*. In S.A. Gelmen&J.P. Byrnes Perspectives on thought and language: Interrelations in development(pp. 225–277). London: Cambridge University Press .
Goldman, A. I. (1992). In defense of the simulation theory. *Mind & Language*, *7*(1–2), 104–119. https://doi.org/10.1111/j.1468-0017.1992.tb00200.x
Gordon, R. M. (1986). Folk psychology as simulation. *Mind and Language*, *1*(2), 158–171.
Grimm, S. R. (2016). How understanding people differs from understanding the natural world. *Philosophical Issues*, *26*(1), 209–225.
Hinds, P. J., Patterson, M., & Pfeffer, J. (2001). Bothered by abstraction: The effect of expertise on knowledge transfer and subsequent novice performance. *Journal of Applied Psychology*, *86*(6), 1232–1243.
Kagan, S. (2001). Thinking about cases. *Social Philosophy and Policy*, *18*(2), 44–63. https://doi.org/10.1017/S0265052500002892
Kappes, H. B., & Morewedge, C. K. (2016). Mental simulation as substitute for experience: Simulations as substitutes. *Social and Personality Psychology Compass*, *10*(7), 405–420. https://doi.org/10.1111/spc3.12257
Keil, F. C. (2006). Explanation and understanding. *Annual Review of Psychology*, *57*, 227–254. https://doi.org/10.1146/annurev.psych.57.102904.190100
Keysar, B., Hayakawa, S. L., & An, S. G. (2012). The foreign-language effect: Thinking in a foreign tongue reduces decision biases. *Psychological Science*, *23*(6), 661–668. https://doi.org/10.1177/0956797611432178
Kinzel, K. (2015). Narrative and evidence. how can case studies from the history of science support claims in the philosophy of science? *Studies in History and Philosophy of Science Part A*, *49*, 48–57.

Kitcher, P. (1982). Implications of incommensurability. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, *1982*, 689–703. Retrieved from JSTOR.

Kitcher, P. (1985). Two approaches to explanation. *The Journal of Philosophy*, *82*(11), 632–639. https://doi.org/10.2307/2026419

Kuhn, T. S. (2012). *The structure of scientific revolutions: 50th Anniversary edition*. Chicago, IL: University of Chicago Press.

Kurby, C. A., & Zacks, J. M. (2013). The activation of modality-specific representations during discourse processing. *Brain and Language*, *126*(3), 338–349. https://doi.org/10.1016/j.bandl.2013.07.003

Lange, M. (2016). *Because without cause: Non-causal explanations in science and mathematics*. Oxford, UK: Oxford University Press.

Lombrozo, T. (2011). The instrumental value of explanations. *Philosophy Compass*, *6*(8), 539–551. https://doi.org/10.1111/j.1747-9991.2011.00413.x

Lombrozo, T. (2012). *Explanation and abductive inference*, https://doi.org/10.1093/oxfordhb/9780199734689.013.0014

Mahr, J., & Csibra, G. (2017). Why do we remember? The communicative function of episodic memory. *The Behavioral and Brain Sciences*, *1–93*, https://doi.org/10.1017/S0140525X17000012

Nichols, S., & Knobe, J. (2007). Moral responsibility and determinism: The cognitive science of folk intuitions. *Nous*, *41*(4), 663–685. https://doi.org/10.1111/j.1468-0068.2007.00666.x

Parikh, N., Ruzic, L., Stewart, G. W., Spreng, R. N., & De Brigard, F. (2018). What if? Neural activity underlying semantic and episodic counterfactual thinking. *NeuroImage*, *178*, 332–345. https://doi.org/10.1016/j.neuroimage.2018.05.053

Paul, L. A. (2014). *Transformative experience*. Oxford, UK: Oxford University Press.

Pennington, N., & Hastie, R. (1988). Explanation-based decision making: Effects of memory structure on judgment. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*(3), 521–533.

Rabinowitz, M., & Gerstel-Friedman, J. (2018). Orienting to see what's important: Learn to ignore the irrelevant. *Quarterly Journal of Experimental Psychology*, *71*(8), 1655–1662.

Richardson, D., & Matlock, T. (2007). The integration of figurative language and static depictions: An eye movement study of fictive motion. *Cognition*, *102*(1), 129–138.

Rosenberg, A. (2018). *How history gets things wrong: The neuroscience of our addiction to stories*. Cambridge, MA: MIT Press.

Sadoski, M. (1983). An exploratory study of the relationships between reported imagery and the comprehension and recall of a story. *Reading Research Quarterly*, *19*(1), 110.

Sadoski, M. (1985). Commentary: The natural use of imagery in story comprehension and recall: Replication and extension. *Reading Research Quarterly*, *20*(5), 658–667. https://doi.org/10.2307/747949

Sadoski, M., Goetz, E. T., & Fritz, J. B. (1993). Impact of concreteness on comprehensibility, interest, and memory for text: Implications for dual coding theory and text design. *Journal of Educational Psychology*, *85*(2), 291–304. https://doi.org/10.1037/0022-0663.85.2.291

Sadoski, M., Goetz, E. T., & Rodriguez, M. (2000). Engaging texts: Effects of concreteness on comprehensibility, interest, and recall in four text types. *Journal of Educational Psychology*, *92*(1), 85–95. https://doi.org/10.1037/0022-0663.92.1.85

Schacter, D. L., & Addis, D. R. (2007). The cognitive neuroscience of constructive memory: Remembering the past and imagining the future. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *362*(1481), 773–786.

Schechtman, M. (2007). Stories, lives, and basic survival: A refinement and defense of the narrative view. *Royal Institute of Philosophy Supplements*, *60*, 155–178. https://doi.org/10.1017/S1358246107000082

Schwartz, D., & Black, T. (1999). Inferences through imagined actions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(1), 116–136.

Scoville, W. B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of Neurology, Neurosurgery & Psychiatry*, *20*, 11–21. https://doi.org/10.1136/jnnp.20.1.11

Sloman, S., Zemla, J. C., Lagnado, D., Bechlivanidis, C., & Hemmatian, B. (in press). Are humans intuitive philosophers? In S. R. Grimm (Ed.), *Varieties of Understanding: New Perspectives from Philosophy, Psychology, and Theology* (pp. 231–250). Oxford, UK: Oxford University Press.

Snow, C. E., & Beals, D. E. (2006). Mealtime talk that supports literacy development. *New Directions for Child and Adolescent Development*, *2006*(111), 51–66. https://doi.org/10.1002/cd.155

Spaulding, S. (2016). Simulation theory. In A. Kind (Ed.), *Routledge handbook of philosophy of imagination* (pp. 262–273). Abingdon, UK: Routledge.

Squire, L. R. (2009). The legacy of patient H.M. for neuroscience. *Neuron*, *61*(1), 6–9. https://doi.org/10.1016/j.neuron.2008.12.023

Strevens, M. (2004). The causal and unification approaches to explanation unified—Causally. *Noûs*, *38*(1), 154–176. https://doi.org/10.1111/j.1468-0068.2004.00466.x

Taylor, C. (1985). *Philosophy and the human sciences*. Cambridge, UK: Cambridge University Press.

Trope, Y., & Liberman, N. (2010). Construal-level theory of psychological distance. *Psychological Review*, *117*(2), 440–463. https://doi.org/10.1037/a0018963

Vossen, P., Caselli, T., & Cybulska, A. (2018). How concrete do we get telling stories? *Topics in Cognitive Science*, *10*(3), 621–640. https://doi.org/10.1111/tops.12366

Walker, C. M., & Lombrozo, T. (2017). Explaining the moral of the story. *Cognition*, *167*, 266–281.

Weatherson, B. (2012). Explanation, idealisation and the Goldilocks problem. *Philosophy and Phenomenological Research*, *84*(2), 461–473.

Wellman, H. M. (2011). Reinvigorating explanations for the study of early cognitive development. *Child Development Perspectives*, *5*(1), 33–38. https://doi.org/10.1111/j.1750-8606.2010.00154.x

Wharton, W. P. (1980). Higher-imagery words and the readability of college history texts. *Journal of Mental Imagery*, *4*(2), 129–147.

Williams, J., & Lombrozo, T. (2010). The role of explanation in discovery and generalization: Evidence from category learning. *Cognitive Science*, *34*(5), 776–806.

Wilkenfeld, D. A., & Lombrozo, T. (2015). Inference to the best explanation (IBE) versus explaining for the best inference (EBI). *Science & Education*, *24*, 1059–1077. https://doi.org/10.1007/s11191-015-9784-4

Zwaan, R. A., Langston, M. C., & Graesser, A. C. (1995). The construction of situation models in narrative comprehension: an event-indexing model. *Psychological Science*, *6*(5), 292–297. https://doi.org/10.1111/j.1467-9280.1995.tb00513.x

Zwaan Rolf, A., Madden, C. J., Yaxley, R. H., & Aveyard, M. E. (2004). Moving words: Dynamic representations in language comprehension. *Cognitive Science*, *28*(4), 611–619.