

2014

# Investigating the Lay and Scientific Norms for Using 'Explanation'

Jonathan Waskan

Ian Harmon

West Virginia University, [ian.harmon@mail.wvu.edu](mailto:ian.harmon@mail.wvu.edu)

Andrew Higgins

Joseph Spino

Follow this and additional works at: [https://researchrepository.wvu.edu/faculty\\_publications](https://researchrepository.wvu.edu/faculty_publications)



Part of the [Philosophy of Science Commons](#)

---

## Digital Commons Citation

Waskan, Jonathan; Harmon, Ian; Higgins, Andrew; and Spino, Joseph, "Investigating the Lay and Scientific Norms for Using 'Explanation'" (2014). *Faculty Scholarship*. 1138.

[https://researchrepository.wvu.edu/faculty\\_publications/1138](https://researchrepository.wvu.edu/faculty_publications/1138)

## Investigating the Lay and Scientific Norms for Using ‘Explanation’

J. Waskan, I. Harmon, A. Higgins, J. Spino  
Psychology of Philosophy Laboratory  
University of Illinois at Urbana-Champaign

### Extended Abstract:

In the mid-20th century, Carl Hempel bucked positivist orthodoxy and proposed that explanations have a legitimate role to play in science. At the same time, in offering his model of explanation, Hempel embraced the positivist tendency to abstract both from facts about human psychology and from the specific contents of explanatory claims. At the broadest level, Hempel believed explanations to be sets of true statements arranged into formally acceptable arguments.

There are still many who would agree that explanations comprise sets of descriptions, or at least representational artifacts – for instance, they might be constructed out of natural or artificial languages, images or diagrams, or computer models). We call this the *representational-artifact* (R-A) view. Many also accept the thesis (which we term *substantive anti-psychologism*) that explanations are not psychological in nature. Recent arguments for the substantive thesis attempt to dissociate explanations from psychological events by noting that some explanations, such as hyper-complex computer models, are so complex that they are incapable of evoking any relevant psychological events (e.g., familiarity, satisfaction, insight, or intelligibility) (see Craver 2006, 2007; Trout 2007). An apparent corollary of the substantive thesis is the methodological thesis that experimental psychology has (apart from supplying case studies of science in general) nothing to teach philosophy about the nature of explanation.<sup>1</sup> It is also commonly claimed that R-A explanations are necessarily accurate (Craver 2007; Hempel 1965; Humphreys 1989; Salmon 1998; Trout 2007).

It is sometimes claimed that at least some explanations are sets of (generally non-representational) objective processes. This is the so-called *ontic* view of explanations. Ontic explanations have also been dissociated from psychological processes (Salmon 1984). As for accuracy, it of course makes little sense to talk of the accuracy of non-representational processes.

There is also a third conception which equates explanations with psychological states. Proponents of this view tend to regard explanations as constituted by psychological representations (Brewer 2001, Keil 2006, Lombrozo & Carey 2006, Thagard (see Simon 1966; Brewer 2001; Vosniadou 2002; Keil 2006; Lombrozo and Carey 2006; Nersessian 2009; Thagard and Litt 2008; Churchland 1989). Experimental research on explanation in this psychological sense commonly focuses on the explanations of children and science-naïve laypeople, explanations which are often to some degree inaccurate.

It may be that all sides are correct in that the noun ‘explanation’ is ambiguous between three senses: an R-A, an ontic, and a psychological sense. One might say, for instance, “There is an aquatic-ape explanation for human furlessness on page thirty seven of *The Naked Ape*.” Here ‘explanation’ refers to a set of descriptive representations which express the possible cause of the

---

<sup>1</sup> The lines are sometimes drawn in terms of ontic, epistemic, and pragmatic conceptions. These, however, are theses about the function of explanations, about what they do rather than about what they are.

target occurrence. Or one might say, “The explanation for combustion is oxidation.” Here ‘explanation’ seems to refer, not to a set of descriptions, but to a set of objective processes that cause the target occurrence. Lastly, one might say, “The pre-verbal child had an explanation for the meowing coming from the closet.” What the child plausibly *has* in this case is a *belief* about the cause of the target occurrence.

To say that ‘explanation’ is ambiguous in this way is to make an empirical claim about the norms of use regarding the term. Such claims are at least tacitly indexed to particular populations, such as the populations of English speaking laypeople or scientists. If it should turn out that there are multiple, distinct senses of ‘explanation,’ this will give rise to the further empirical question of what distinguishes the three, both semantically and evidentially. Knowing this, we will be in a better position to study the role of the entities referred to by the various senses of ‘explanation’ in our lay and scientific lives. Like Hempel, we maintain that it is essential to understand what explanations are and what role they play if we are to understand how science works and what makes it such an exemplary epistemic exercise. It may be that, once we have a deeper grasp of the reality, we may wish to prescribe modifications or precisifications to the terminology that we use to describe it. Here too, we believe, one must know something about actual usage.

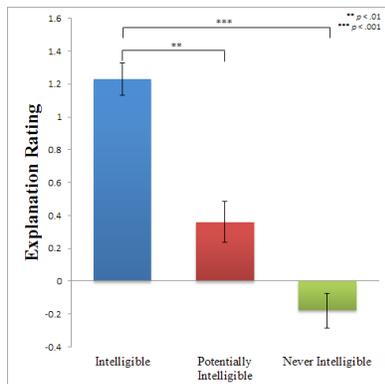
One common way in which philosophers of science go about answering empirical questions about how people (laypeople or scientists) use the term ‘explanation’ is to consult their own judgments about cases. For instance, one might imagine a hyper-complex model of a physical process. One might imagine further that it has any number of positive theoretical virtues but that nobody will ever fully comprehend it. One might then judge that such a model is, or could well be, an obvious or clear-cut case of explanation. This is to say, we think, that some target population (one which presumably includes scientists) would also classify such a model as an explanation. From this supposition, again, one might argue that an artifact’s ability to engender psychological states (take your pick) in humans is irrelevant to their standing as explanations. Practiced in isolation, this method is clearly inadequate, as it presupposes that the philosophers in question have mastered the relevant scientific and linguistic norms and that their judgments are not merely issuing from the very theories they are meant to prove (Cummins 1998).

Another strategy, which is a staple in the philosophy of science, is to look to the historical record for evidence that supports or undermines a given theory. Philosophers often look to the writings, graphs, and diagrams that are the products of scientific research. To do this effectively, one must be careful that one’s sample is adequately large, representative, and unbiased in its selection (Thagard 2012). These requirements are seldom met in practice. For one thing, research of this sort often focuses on the most noteworthy and revolutionary developments in science. Either that, or philosophers pick from a small number of particular developments of which they have in-depth knowledge. Like the previous strategy, this one is under-constrained by information about the actual facts regarding usage.

Generally speaking, we think that in this area, philosophy of science would do well to expand its arsenal. On this view, the ideal theory of the facts about science will be one that has converging evidence from a variety of sources.

Recent work by Overton (2013) illustrates one possible source of additional data (and constraint). Overton undertakes the bibliometric investigation of ‘explain’ and its cognates in recent scientific articles. Although he only considers the semantics of explanation-talk in a small set of case studies, the basic techniques employed could be expanded to take in a much wider range of cases.

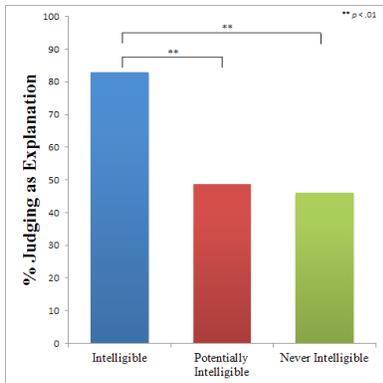
One might also gain relevant data by applying the methods of experimental and social psychology. Studies of this sort have already been conducted in order to discern how the use of terms such as ‘gene’ and ‘innate’ varies between laypeople and scientists and how it can also vary within these two populations (Stotz and Griffiths 2008). We pursue a similar line of investigation with regard to how ‘explanation’ is employed by laypeople and scientists. In earlier studies (Braverman et al. 2012; Waskan et al. under review) we focused exclusively on the R-A conception. We found that, contrary to what many philosophers of science maintain, both laypeople and scientists regard R-A explanations as inherently psychological. In particular, they seem to regard it as a crucial, or at least central, component of explanations that they have actually rendered the target intelligible (i.e., supplied understanding of how or why-possibly an event occurred) to someone. The mere *capacity* to do so does not suffice. This finding held for laypeople under two very different paradigms. The first (Fig. 1) involved traditional brief vignettes (~200 words) followed by ratings questions.



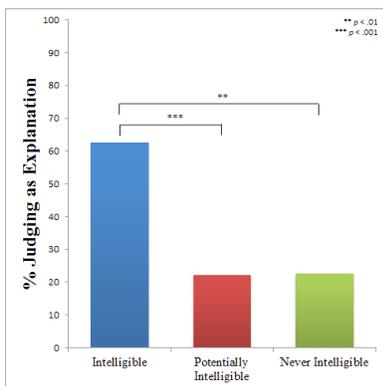
**Figure 1.** Mean ‘constitutes an explanation’ ratings by condition.

The second utilized a lengthy article (~1000 words) that was said to have been published in a well-known science magazine (in fact, only part had been published). Rather than asking for an explicit rating, which we feel can lead participants to engage in amateur philosophizing, the explanation probe was disguised as part of a larger memory task. The task involved determining whether or not particular sentences seemed likely to be true based upon the material in the article. Rather than recruiting episodic memories and explicit deliberation, such ‘semantic integration’ tasks recruit semantic memories of the meaning conveyed (Powell, Horne, and Pinillos forthcoming).

Using this second paradigm, we continued to find that laypeople (Fig. 2) place great value on actual intelligibility. This experiment was repeated using practicing scientists as subjects and the results were comparable (Fig. 3).



**Figure 2.** Percentage of MTurk participants recalling that an explanation was provided.



**Figure 3.** Percentage of professional scientists recalling that an explanation was provided.

In our latest study, we have undertaken to compare the application conditions of the R-A and ontic senses of ‘explanation’ for both laypeople and scientists. It seemed to us at least possible that intelligibility is also important to the ontic conception – for instance, it could be that an objective process is only regarded as an explanation if someone can possibly understand how it brings about the target occurrence. We thus compared memory for the presence of an R-A explanation and for an ontic explanation under two conditions:

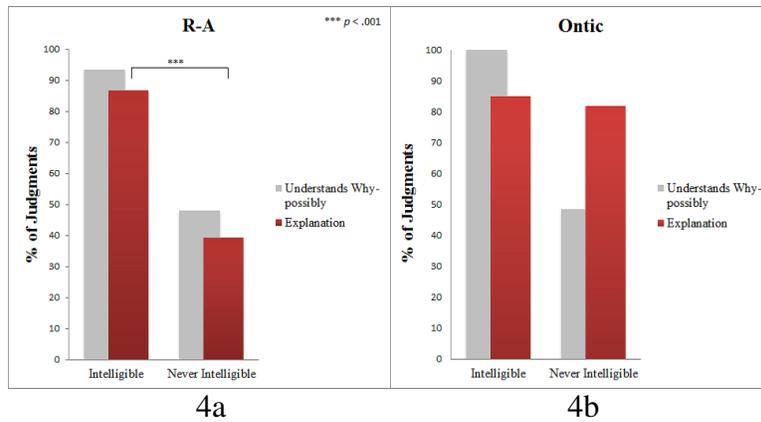
Intelligible – materials (viz., a hyper-complex computer model and a published article regarding it) that are packed with standard theoretical virtues, and that have actually rendered a target occurrence intelligible to a particular scientist (Brown).

Never Intelligible – materials (viz., a hyper-complex computer model and a published article regarding it) that are packed with standard theoretical virtues, and that are (because of their complexity) incapable of rendering the target occurrence intelligible to anyone.

We also included a lengthy distractor (a different article about neuroscience) so as to heighten reliance upon semantic memory. As part of the memory task, we included a measure of whether or not participants remember that the materials enabled Brown to understand how the target occurrence comes about (labeled ‘Understands how-possibly’ in Figure 4).

We found (as before) that judgments that an R-A explanation was provided varies between the two conditions (Figure 4). Notice as well that judgments regarding whether or not Brown understands how-possibly varies to a similar extent across the two conditions. Understanding how-possibly judgments vary to about the same extent when memory for ontic explanation is being tested. Ontic explanation judgments, however, are unaffected by manipulating

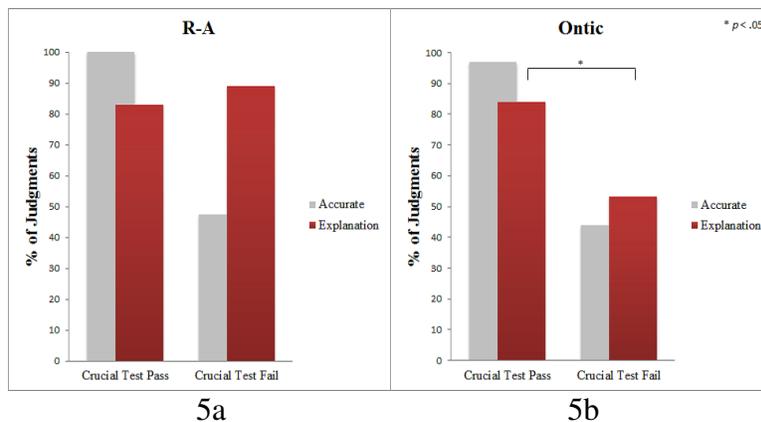
intelligibility. This suggests that intelligibility is not taken to be required for ontic explanation. This is important because intelligibility may be the most likely candidate for a psychological event that people would regard as crucial, or central, for ontic explanation.



**Figure 4.** Percentage of participants judging under Intelligible and Never Intelligible conditions that the story describes an R-A (4a) or ontic (4b) explanation.

In short, while R-A explanation judgments are affected by the presence or absence of intelligibility, ontic explanation judgments are not. These judgments, we believe, are a good indicator of the semantic content conveyed to participants by the article.

We wondered if these two forms of explanation might come apart in other ways as well, such as by manipulating the sense that the materials in question are accurate. On most views, accuracy is essential for R-A explanation, so manipulating accuracy should alter R-A explanation judgments. Also, insofar as the materials are inaccurate, the physical process they designate will not be regarded as the (ontic) explanation for the occurrence. Thus manipulating accuracy should also alter ontic explanation judgments. To alter accuracy judgments, we varied whether or not the materials supplied by Brown survive a crucial test. We include an added measure of whether or not participants recall, after the lengthy distractor period, that the materials Brown supplied were accurate (labeled ‘Accurate’ in Figure 5).



**Figure 5.** Percentage of participants judging under the Crucial-pass and Crucial-fail conditions that the story describes an R-A (5a) or ontic (5b) explanation.

What we found was that the crucial-test manipulation had a powerful impact on lay participants' judgments regarding the accuracy of the particular materials and regarding whether or not the process they designate constitute an ontic explanation (Figure 5). Importantly, this manipulation did not significantly affect judgments about whether or not the materials constitute an R-A explanation. This again runs counter to the conclusions reached utilizing more traditional philosophical methods (i.e., that accuracy is necessary for R-A explanation).

Currently we are gathering data on scientists to determine whether accuracy and intelligibility affect their judgments about R-A and ontic explanations. From the data collected thus far, it appears that scientists' judgments mirror those of lay participants in the above-mentioned respects.<sup>2</sup>

Much more research is needed in order to determine whether or not there is a third, psychological sense of 'explanation' in use, which is distinct from the R-A and ontic senses, and what sorts of factors influence psychological explanation judgments. What the above line of research illustrates clearly, however, is that there is an alternative means of studying the semantic and epistemic application criteria of 'explanation'. This method is distinct from the intuition-based, case-study based, and representative sample-based methods for studying the norms in use with regard to key scientific terms. We have seen that the results generated through the use of this new method may come into conflict with the results of those other methods that do not involve eliciting scientists' judgments directly. Clearly something has to give. We leave it as a matter open to discussion as to how such conflicts are best resolved.

- Braverman, M., J. Clevenger, I. Harmon, A. Higgins, Z. Horne, J. Spino, and J. Waskan. 2012. Intelligibility is necessary for explanation but accuracy may not be. Paper read at Proceedings of the Thirty-Fourth Annual Conference of the Cognitive Science Society.
- Brewer, W.F. 2001. Models in science and mental models in scientists and nonscientists. *Mind & Society* 2 (2):33-48.
- Churchland, P.M. 1989. *A neurocomputational perspective: The nature of mind and the structure of science*. Cambridge, MA: The MIT Press.
- Craver, C.F. 2006. When mechanistic models explain. *Synthese* 153 (3):355-376.
- Craver, C.F. 2007. *Explaining the brain*. New York: Oxford University Press.
- Cummins, Robert. 1998. Reflections on reflective equilibrium. In *The Role of Intuition in Philosophy*, edited by W. a. M. D. Ramsey. New York: Roman and Littlefield.
- Hempel, C.G. 1965. *Aspects of scientific explanation and other essays in the philosophy of science*. New York: The Free Press.
- Humphreys, P. 1989. Scientific explanation: The causes, some of the Causes, and nothing but the causes. *Minnesota studies in the philosophy of science* 13:283-306.
- Keil, F.C. 2006. Explanation and understanding. *Annual Review of Psychology* 57:227-254.
- Lombrozo, Tania, and Susan Carey. 2006. Functional explanation and the function of explanation. *Cognition* 99 (2):167-204.
- Nersessian, N.J. 2009. How do engineering scientists think? Model-based simulation in biomedical engineering research laboratories. *Topics in Cognitive Science* 1:730-757.
- Overton, J. 2013. "Explain" in scientific discourse. *Synthese* 190:1383-1405.

---

<sup>2</sup> It is too early to say with certainty how the data will look after all participants have responded, but this data will be in early May.

- Powell, D., Z. Horne, and N. Pinillos. forthcoming. Semantic Integration as a method for investigating concepts. In *Advances in Experimental Epistemology*, edited by J. Beebe. New York: Continuum Press.
- Salmon, W. 1984. *Scientific explanation and the causal structure of the world*. Princeton, NJ: Princeton University Press.
- Salmon, W. 1998. *Causality and explanation*. New York: Oxford University Press.
- Simon, H.A. 1966. Thinking by computers. In *Mind and Cosmos: Essays in Contemporary Science and Philosophy*, edited by R. G. Colodny. Pittsburgh: University of Pittsburgh Press.
- Stotz, K., and P.E. Griffiths. 2008. Experimental philosophy of science. *Philosophy Compass* 3 (3):507-521.
- Thagard, P. 2012. *The Cognitive Science of Science: Explanation, Discovery, and Conceptual Change*. The MIT Press.
- Thagard, P., and A. Litt. 2008. Models of scientific explanation. In *The Cambridge Handbook of Computational Cognitive Modeling*, edited by R. Sun. Cambridge, MA: Cambridge University Press.
- Trout, JD. 2007. The psychology of scientific explanation. *Philosophy Compass* 2 (3):564-591.
- Vosniadou, S. 2002. Mental models in conceptual development. In *Model-based reasoning: Science, technology, values*, edited by L. Magnani and N. J. Nersessian. Berlin: Springer.
- Waskan, J., J. Clevenger, I. Harmon, Z. Horne, and J. Spino. under review. Explanatory anti-psychologism overturned by lay and scientific case classifications.