

Surrogates for Theory

March 08, 2011

Science walks forward on two feet, namely theory and experiment ... Sometimes it is one foot that is put forward first, sometimes the other, but continuous progress is only made by the use of both.

Robert A. Millikan, Nobel Lecture 1924

Psychologists treat other people's theories like toothbrushes – no self-respecting person wants to use anyone else's.

Walter Mischel, Observer 2008

Will 21st-century psychology walk forward on two feet, or hobble on one? Do we teach our graduate students that the goal of science is advancement of explanation through theory and that experimentation is the tool to ensure we actually do advance? It appears to me much of psychology is split into two camps. One is theory-guided, designing computational or analytic models of cognitive processes and testing predictions against alternative accounts, a process that allows theories to be improved. The other camp's philosophy is theoretical minimalism: It makes no noticeable attempt to design a theory. In my view, this second camp limps along on one leg and a crutch: the one leg is experimentation, the crutch is a *surrogate* for theory. *

The problem is not that a majority of researchers would say that theory is irrelevant; the problem is that almost anything passes as a theory. I have encountered three species of surrogates: one-word explanations, circular restatements, and lists of dichotomies. If you know of more, let me know. What distinguishes these surrogates from genuine theory is that they can account for almost everything after the fact but cannot make new predictions. They are useful because they are so vague that they can't be wrong. To keep this article brief yet remain thought-provoking, I have had to caricature, but not too much.

One-Word Explanations

It works like this: To explain a phenomenon, pick an English noun that is sufficiently broad in meaning, but don't develop it into a model by specifying an underlying mechanism. The result is a label with all the virtues of a Rorschach inkblot: One can see lots of things in it. For instance, for over a century, memory theories have distinguished between recall, recognition, recency, and frequency, and provided quantitative models that make predictions. All that sophistication has been lost in some areas of psychology and been replaced by one single word: availability. At times this term denotes the "number" of instances that come to mind, at times the "ease" with which instances come to mind, and at yet other times something else: recency, salience, memorability, vividness, and more. When ease and number were actually measured, they were found to differ and, most important, to *not* correlate with the frequency judgments that the term *availability* purportedly explains (Sedlmeier et al., 1998). The few attempts to flesh out this label into a testable model met with little enthusiasm.

One-word explanations are so perfectly flexible that the same word can be used to account for a phenomenon and its opposite, that is, for A and not-A. Consider the gambler's fallacy: After a series of n reds on the roulette table, the expectation of another red *decreases*. This intuition was explained by people's reliance on "representativeness." Next consider the hot-hand fallacy, which is the opposite intuition: After a basketball player scores a series of n hits, the expectation of another hit *increases*. This belief was also attributed to representativeness.

No theory should predict both a phenomenon and its contrary. But a label can do this by constantly changing its meaning: To account for the gambler's fallacy, the term alludes to a higher similarity between the series of $n + 1$ outcomes and the underlying chance process, whereas to account for the hot-hand fallacy, it alludes to a similarity between a series of n and observation number $n + 1$. As with availability, we witness an actual decline in theoretical precision. The seminal theories such as Shepard's symmetric and Tversky's asymmetric models of similarity have been largely replaced by one word: representativeness.

Circular Restatements

Recall Molière's parody of the Aristotelian doctrine of substantial forms: Why does opium make you sleepy? Because of its dormative properties. Here is how circular restatements work: The observation that A influenced B is "explained" by saying that A had the propensity to influence B. While the use of one-word explanations still distinguishes between the observation and its explanation, circular restatements can be defined by the absence of this distinction. The explanation *is* the phenomenon, merely couched in slightly different terms.

Thus, when researchers observed that people are influenced both by the logical form of a syllogism and the believability of its conclusion, this belief-bias effect was explained by the operation of two reasoning systems, one logical and the other based on prior belief. When participants were given the Wason selection task and a rule "if there is an A on one side, then there is [not] a 3 on the other," the finding was that many checked the "A" and the "3" card, whether the "not" was present or not. This observation was explained by a "matching bias," which means that the participants consider the two cards mentioned in the rule as relevant information. These explanations are as close to circular restatements as they can be. When two different external representations of the same information (framing) resulted in different performances, the restated explanation was that one of these makes the solution more "transparent" or "salient." That a representation makes a solution transparent is hardly an explanation, however, but rather what needs to be explained.

One can also find circular restatements in the intersection of psychology and economics. Why do many people divide money equally in the ultimatum game, violating the predictions of game theory? Circular restatement responds that they have an inequality aversion. Finally, the aforementioned *availability heuristic* was invoked to explain various cognitive biases, but recently the term *availability bias* has come to be used interchangeably. Here, the same term refers to the phenomenon and its explanation. Like one-word explanations, circular restatements are immune to improvement. They can't be wrong.

Lists of Dichotomies

Torn between dismay and satisfaction concerning the then-state of research on information processing,

Allen Newell (1973) wrote a commentary entitled “You can’t play 20 questions with nature and win.” What distressed Newell was the fact that when behavior is explained in terms of dichotomies—nature versus nurture, serial versus parallel, grammars versus associations, and so on—“clarity is never achieved” and “matters simply become muddier and muddier as we go down through time.” There is nothing wrong with dichotomies per se; what concerned Newell are situations in which theoretical thinking gets stuck in binary oppositions beyond which it never progresses.

Thirty-odd years later, some areas of social and cognitive psychology are once again stuck with a yin-yang list of dichotomies. Here is how it works: Construct a list of general oppositions and call the one side of the list System 1 and the other System 2. Do not construct models of the processes involved and ignore existing models. Almost every behavior can be attributed post hoc to one of the two systems, but new predictions can rarely if ever be derived. Again, the theoretical development is backwards; we have devolved from elaborated process models to a mere list of general oppositions.

Consider one dichotomy that appears in most of these lists: associations versus rules. What does this opposition mean? It has been linked to Smolensky’s distinction between an intuitive processor and a conscious rule interpreter, to Hinton’s distinction between intuitive and rational processing, to Schneider and Shiffrin’s distinction between automatic and controlled processing, to Evans’s distinction between a perceptually-based matching process and a linguistic-logical process, and to Freud’s distinction between primary and secondary processes. Yet these oppositions are not the same (Gigerenzer & Regier, 1996), and much of the clarity in the cited theories is lost in the opaque dichotomy.

System 1 is often called “heuristic” without saying what the heuristic processes are. There exists a rich body of formal models of heuristics — from Simon to Tversky to Payne, Bettman, and Johnson to my own work — that has been replaced by a single black box. System 1 is also called “unconscious”, although we know that doctors, managers, burglars, and other experts often use heuristics deliberately. Finally, System 2 is often called “logical” or “reflective,” without these logical processes being specified. There is a large and refined literature on various logical, Bayesian, Neyman-Pearsonian, and sequential decision theories that do not speak with one voice and can lead to divergent predictions (Oaksford & Chater, 1994). To reduce all of this knowledge to a few dichotomies is a step backwards.

Theory Integration

Developing a theory, even one that is backed by models of mechanisms and fertile in testable predictions, is not enough. The toothbrush problem remains: If all cling to their personal theories, psychology will never become a powerful enterprise. In fact, much of the theoretical landscape in psychology resembles a patchwork of small territories. Isn’t our goal theory integration, constructing a larger unified territory out of small fiefdoms? While in physics and biology the unification of theories, such as evolutionary theory and genetics, is a widely shared goal, this aspiration is barely visible in psychology. Integration can begin by taking two theoretical systems and showing how connecting them can lead to new insights. For instance, relating the ACT-R cognitive architecture program to the Adaptive Toolbox program has led to the discovery of how systematic forgetting leads to better heuristic inferences (Schooler & Hertwig, 2005).

I see two steps for making progress towards a cumulative science. The first is to replace surrogates by genuine theories (to walk forward on two feet), and the second is to begin theory integration (to walk

together). Let me end with a story. Some years ago I spent a day and a night in a magnificent library reading through issues of the *Journal of Experimental Psychology* from the 1920s and 1930s. Much of today's research pales in comparison to the diversity of methods and statistics, the detailed reporting of single-case data rather than mere averages, and the careful selection of trained subjects on display in many of the articles published way back then. What depressed me was that nearly all of this meticulous work has been forgotten. Most of it involved collecting data without substantive theory. Data without theory is like a baby without a parent: Its life expectancy is low. Are these the kind of babies we want to produce?

* I will give specific examples of surrogates in the following, but omit the names of the authors, since at issue are not individuals but a research philosophy that hurts psychology. It is easy to find these and many more examples in the literature.