Published in: *Theory & Psychology, 8* (2), 1998, 195–204. http://tap.sagepub.com/

Surrogates for Theories

Gerd Gigerenzer
Max Planck Institute for Human Development

Abstract. I first discuss several strategies that serve as surrogates for theories in psychology: one-word explanation, redescription, drawing vague dichotomies, and data fitting. I then identify two conventions that allow these surrogates to flourish and multiply: null hypothesis testing, which makes precise hypotheses irrelevant, and the isolation of research in different disciplines, which prevents the exchange of positive metaphors between fields.

Key Words: explanation, null hypothesis testing, redescription, surrogates, theory construction.

I like conference dinners. At such a dinner several years ago, I was crammed in with four graduate students and four professors around a table laden with Chinese food. The graduate students were eager to learn first-hand how to complete a dissertation and begin a research career, and the professors were keen to give advice. With authority, one colleague advised them: "Don't think big. Just do four or five experiments, clip them together, and hand them in." The graduate students nodded gratefully. They continued to nod when I added: "Don't follow this advice unless you are mediocre or unimaginative. Try to think in a deep, bold, and precise way. Take risks and be courageous." What a dilemma. How could these students apply these contradictory pieces of advice?

Based on an analysis of articles in two major social psychology journals, the *Journal of Personality and Social Psychology* and the *Journal of Experimental Social Psychology,* Wallach and Wallach (1994, 1998) concluded that the theoretical argument in almost half of the studies reported borders on tautology. If an argument is a "near-tautology," there is no point in spending time and money to try to experimentally confirm it. "Don't think big" seems to be a prescription followed by many professional researchers, not merely conservative advice for graduate students. Complaints about the lack of serious theory in social psychology have been voiced before (e.g., Fiedler, 1991, 1996). Atheoretical research is not specific to social psychology, however, although some parts of psychology do better than others (Brandtstädter, 1987).

In this article, I will address two questions. What are the surrogates for theory in psychology? What institutional forces perpetuate reliance on these surrogates? This essay is not intended to be exhaustive, only illustrative. The examples I use are drawn from the best work in the areas discussed: the psychology of reasoning, judgment, and decision making.

Surrogates

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 identify four species of surrogates for theory: one-word explanations, redescription, muddy dichotomies and data fitting. There are other species, such as Wallach and Wallach's (1994, 1998) near-tautologies (whose banality and lack of imag-

ination are as striking as their circularity). What distinguishes these surrogates from genuine theory is that they are vague, imprecise, and/or practically unfalsifiable, that they often boil down to common sense and lack boldness and surprise.

One-Word Explanations

The first species of theory surrogate is the one-word explanation. Such a word is a noun, broad in its meaning and chosen to relate to the phenomenon. At the same time, it specifies no underlying mechanism or theoretical structure. The one-word explanation is a label with the virtue of a Rorschach inkblot: a researcher can read into it whatever he or she wishes to see.

Examples of one-word explanations are *representativeness, availability,* and *anchoring and adjustment,* which are treated as the cognitive heuristics people use to make judgments and decisions. These terms supposedly explain "cognitive illusions," such as base-rate neglect. These "explanations" figure prominently in current textbooks in cognitive psychology, social psychology, and decision making. It is understandable that when these three terms were first proposed as cognitive processes in the early 1970s, they were only loosely characterized (Tversky & Kahneman, 1974). Yet 25 years and many experiments later, these three "heuristics" remain vague and undefined, unspecified both with respect to the antecedent conditions that elicit (or suppress) them and also to the cognitive processes that underlie them (Gigerenzer, 1996). I fear that in another 25 years we will still be stuck with plausible yet nebulous proposals of the same type: that judgments of probability or frequency are sometimes influenced by what is similar (representativeness), comes easily to mind (availability), and comes first (anchoring).

The problem with these heuristics is that, post hoc, at least one of them can be fitted to almost any experimental result. For example, base-rate neglect is commonly attributed to representativeness. But the opposite result, overweighting of base rates ("conservatism"), is just as easily "explained" by invoking anchoring (on the base rate) and adjustment. One-word explanations derive their seductive power from the fact that almost every observation can be called upon as an example.

One might think that researchers rely on such one-word explanations because they lack precise models. But this is not the case. For instance, there are several precise definitions of "similarity," such as Euclidean distance, the city-block metric and various measures of feature overlap, including Tversky's (1977) model (Shepard, 1962, 1974). Furthermore, recent attempts to define the word "representativeness" have met with little attention (e.g., Smith & Osherson, 1989). Proponents of one-word explanations continue to defend undefined terms. For instance, it has been recently argued that representativeness "can be assessed experimentally; hence it need not be defined a priori" (Kahneman & Tversky, 1996, p. 585). The term "availability" is similarly vague. Sometimes it denotes the "number" of instances that come to mind, sometimes the "ease" with which instances come to mind, and at still other times it means something else. Again, some researchers have attempted to pin down the meaning of the term (e.g., Fiedler, 1983, 1991; Wänke, Schwarz, & Bless, 1995), but with little effect on the thinking of others in the field. There is also a long tradition of fairly precise conceptions of "anchoring," such as in Helson's adapation level and Parducci's range-frequency theories. But the seductive power of one-word explanations seems to have caused collective amnesia. The strange reluctance of many researchers of reasoning, judgment, and decision making to specify precise and falsifiable process models and to work out the relationship between cognitive heuristics has been pointed out repeatedly (e.g., Einhorn & Hogarth, 1981; Shanteau, 1989; Wallsten, 1983).

But one-word explanations have great advantages. As long as they are plausible and remain unspecified, they are hard to falsify. And if one has three to choose from—such as representativeness, availability, and anchoring and adjustment—at least one of them can "account" post hoc for almost any phenomenon. The near-omnipotence of one-word explanations, however, does not foster theory development.

Redescription

Recall Molière's parody of the Aristotelian doctrine of substantial forms: Why does opium make you sleepy? Because of its dormative properties. Redescription has a long tradition in trait psychology, for instance, when an aggressive behavior is attributed to an aggressive disposition, or intelligent behavior to high intelligence. But redescription in psychology is not limited to attributing behaviors to traits and other essences.

Research on reasoning and judgment and decision making is another field in which redescription flourishes as a surrogate for theoretical ideas. Hallmarks of redescription in these areas are words such as "transparent," "relevant," and "salient" (which is not to say that every use of these terms implies redescription). For instance, one important issue in problem solving is how the external representation of a problem—whether it is represented in the form of text, figures, probabilities, frequencies, and so on—influences performance. If an effect is found, the question arises: Why does this type of problem representation elicit better performance? Redescription creeps in when researchers propose that the effect was obtained because "the correct answer is made transparent by the representation," because of "a salient cue that makes the correct answer obvious," or because "the problem is now simpler" (see Gigerenzer, 1996; Gigerenzer & Murray, 1987, pp. 159–162; Sahlin, 1991). That a representation makes a problem "simpler" or its answer "obvious" is not an explanation, but rather what needs to be explained.

Muddy Dichotomies

Torn between being distressed and content with the state of research on information processing, Allen Newell (1973) entitled a commentary "You Can't Play 20 Questions with Nature and Win." What distressed Newell was that when behavior is explained in terms of dichotomies—nature vs. nurture, serial vs. parallel, grammars vs. associations, and so on—"clarity is never achieved" and "matters simply become muddier and muddier as we go down through time" (pp. 288–289). There is nothing wrong with making distinctions in terms of dichotomies per se; what concerned Newell were situations in which theoretical thinking gets stuck in binary oppositions beyond which it never seems to move.

Let us consider a case in which false dichotomies have hindered precise theorizing. Some arguments against evolutionary psychology are based on the presumed dichotomy between biology and culture, or genes and environment (Tooby & Cosmides, 1992). One such argument goes: Since cognition is bound to culture, evolution must be irrelevant. But biology and culture are not opposites. For instance, our ability to cooperate with conspecifics to whom we are genetically unrelated—which distinguishes us humans from most other species—is based on mechanisms of both biological and cultural origin. Simply to ask about the relative importance of each in terms of explained variance, such as that 80 percent of intelligence is genetically inherited, is, however,

not always an interesting question. The real theoretical question concerns the mechanism that combines what is termed the "biological" and the "cultural." For biologists, the nature/nurture or biological/cultural dichotomy is a non-starter: genes are influenced by their environment, which can include other genes.

Cognitive psychology is also muddied by vague dichotomies. For instance, a popular opposition is between associations and rules. Sloman (1996) has linked this dichotomy to Smolensky's distinction between an intuitive processor and a conscious rule interpreter, Hinton's distinction between intuitive and rational processing, Schneider and Schiffrin's distinction between automatic and controlled processing, Evans' distinction between a perceptually based matching process and a linguistic-logical process, and Freud's distinction between primary and secondary processes. The problem is that these distinctions are not all the same (Gigerenzer & Regier, 1996). Collecting more and more binary oppositions does not necessarily enhance clarity. Dichotomies can be an important first step, but they cannot substitute for theories of cognitive processes.

Data Fitting

There are other surrogates for theories in psychology, one of which is the use of powerful mathematical tools for data fitting in the absence of theoretical underpinnings. Psychologists have historically embraced such new tools, which they then propose as new theories. When factor analysis became a common tool for data processing in psychological research, humans were modeled as a bundle of personality factors. When multidimensional scaling came along in the 1960s and 1970s, human categorization and other mental processes were proposed to be based on distances between points in multidimensional space. More recently, the advent of the serial computer was followed by that of neural networks as a model of cognitive function (Gigerenzer, 1981, 1992). There is nothing wrong with using these mathematical tools per se. The important point with respect to surrogate theories is whether the tool is used for modeling or for data fitting (this is itself a false dichotomy, there being a continuum between these poles). Charles Spearman originally designed factor analysis as a theory of intelligence, but (in the form of principal component analysis) it ended up as a fitting tool for all kinds of psychological phenomena. Likewise, Roger Shepard (e.g., 1962, 1974) interpreted the various Minkowski metrics that can be used in multidimensional scaling as psychological theories of similarity, such as in color perception, but multidimensional scaling ended up as a largely atheoretical tool for fitting any similarity data, with the Euclidean metric as a conventional routine. Similarly, neural networks can be used as constrained or structured networks into which theoretical, domain-specific assumptions are built (e.g., Regier, 1996), but many applications of neural networks to modeling psychological phenomena seem to amount to data fitting with numerous free parameters. Neural networks with hidden units and other free parameters can be too powerful to be meaningful—in the sense that they can fit different types of results that were generated with different process models (Massaro, 1988).

In general, mathematical structures can be used to test theories (with parameters determined by theoretical considerations, e.g., the metric in multidimensional scaling) or as a fitting tool (with parameters chosen post hoc so as to maximize the fit). Fitting per se is not objectionable. The danger is that enthusiasm for a mathematical tool can lead one to get stuck in data fitting, and to use a good fit as a surrogate for a theory.

What Institutional Forces Support Surrogates for Theories?

There is one obvious reason why surrogates for theories come to mind more quickly than real theories: demonstrating how a one-word explanation, a redescription, a dichotomy or an excercise in data fitting "explains" a phenomenon demands less mental strain than developing a bold and precise theory. It takes imagination to conceive the idea that heat is caused by motion, but only little mental effort to propose that heat is caused by specific particles that have the propensity to be hot. In what follows, I identify two institutions that may maintain (rather than cause) the abundant use of surrogates for theories in some areas of psychology.

The Institutionalization of Null Hypothesis Testing

In recent years, more and more scholars have argued against the ritual of null hypothesis testing, which was institutionalized in psychology around 1955 (for a discussion, see Gigerenzer et al., 1989, ch. 3 and 6). Many other scholars have responded with passionate defenses of it. So far the debate focuses on issues such as whether one should replace significance testing by confidence intervals, effect sizes, or something else. In my view, however, the single most important issue is that institutionalized null hypothesis testing allows surrogates for theories to flourish (Gigerenzer, 1993). To switch to confidence intervals will not necessarily reverse this trend.

Null hypothesis testing provides researchers with no incentive to specify either their own research hypotheses or competing hypotheses. The ritual is to test one's unspecified hypothesis against "chance," that is, against the null hypothesis that postulates "no difference between the means of two populations" or "zero correlation." As Danziger (1990) has shown, the origin of this practice is in parapsychology and education, where the interest was not in testing positive theories but detecting effects greater than those of chance. The problem lies not in statistical testing per se, but in a specific statistical method that became institutionalized. If psychologists had adopted statistical methods that test two or more well-specified hypotheses against one another (such as Neyman-Pearson or Bayesian statistics), then they would have been forced to formulate precise hypotheses.

As long as there is an institutionalized methodology that does not encourage researchers to specify their hypotheses, there is little incentive to think hard and develop theories from which such hypotheses could be derived. Accordingly, the voluminous instruction manual from which graduate students and researchers learn how to write an article, the *APA Publication Manual,* devotes many pages to methodology but not to theory. From this students learn not that hypotheses and theories should be bold, surprising, and precisely stated, but that their business is to test null hypotheses. To perform this ritual, mere surrogates for thinking big are sufficient. The result has been called "null science" (Bower, 1997).

Disciplinary Isolation

Over the course of the 20th century, academic psychology has become more and more compartmentalized into subdisciplines such as social psychology, cognitive psychology, developmental psychology, and so on. Each subdiscipline has its own journals, reviewers, and grant programs, and one can have a career in one of them without ever reading the journals of neighboring subdisciplines. In addition, job searches are often organized according to these categories. This ter-

ritorial organization of psychology discourages researchers from engaging with psychological knowledge and colleagues outside of their territory, not to mention with other disciplines. As Jerry Fodor (1995) put it:

> Unfortunately, cognitive psychology as people are trained to practice it, at least in this country, has been traditionally committed to methodological empiricism and to disciplinary isolationism, in which it was, for example, perfectly possible to study language without knowing anything about linguistics. (pp. 85–86)

This isolationism is by no means restricted to the study of language. For instance, the experimental study of logical thinking in arguably the most researched problem, the Wason selection task, has been carried out with little reference to modern logic, and the study of statistical reasoning has been conducted with little attention to the relevant issues in statistics (see Gigerenzer, 1994; Oaksford & Chater, 1994).

Intellectual inbreeding can block the flow of positive metaphors from one discipline to another. Neither disciplines nor subdisciplines are natural categories. Interdisciplinary exchange has fueled the development of some of the most influential new metaphors and theories in the sciences, such as when Ludwig Bolzmann and James Clerk Maxwell developed statistical mechanics by borrowing from sociology. Bolzmann and Maxwell modelled the behavior of gas molecules after the behavior of humans as Adolphe Quetelet had portrayed it: erratic and unpredictable at the individual level, but exhibiting orderly statistical laws at the level of collectives (Gigerenzer et al., 1989, ch. 2). Territorial science, in contrast, blocks the flow of metaphors and the development of new theories. Distrust and disinterest in anything outside one's subdiscipline supports surrogates for theory.

Data Without Theory

In this article, I have specified four surrogates for theory and two possible institutional reasons why some of these surrogates flourish like weeds. These two reasons certainly cannot explain the whole story.

Several years ago, I spent a day and a night in a library reading through issues of the *Journal of Experimental Psychology* from the 1920s and 1930s. This was professionally a most depressing experience. Not because these articles were methodologically mediocre. On the contrary, many of them make today's research pale in comparison to their diversity of methods and statistics, their detailed reporting of single-case data rather than mere averages, and their careful selection of trained subjects. And many topics—such as the influence of the gender of the experimenter on the performance of the participants—were of interest then as now. What depressed me was that almost all of this work is forgotten; it does not seem to have left a trace in the collective memory of our profession. It struck me that most of it involved collecting data without substantive theory. Data without theory are like a baby without a parent: their life expectancy is low.

References

Bower, B. (1997). Null science. *Science News, 151,* 356–357.

Brandtstädter, J. (1987). On certainty and universality in human development: Developmental psychology between apriorism and empiricism. In M. Chapman & R. A. Dixon (Eds.), *Meaning and the growth of understanding: Wittgenstein's significance for developmental psychology* (pp. 69–84). Berlin: Springer.

Danziger, K. (1990). *Constructing the subject.* Cambridge, England: Cambridge University Press.
Einhorn, H. J., & Hogarth, R. M. (1981). Behavioral decision theory: Processes of judgment and choice. *Annual Review of Psychology, 32,* 53–88.
Fiedler, K. (1983). On the testability of the availability heuristic. In R. W. Scholz (Ed.), *Decision making under uncertainty* (pp. 109–119). Amsterdam: North-Holland.
Fiedler, K. (1991). Heuristics and biases in theory formation. *Theory & Psychology, 1,* 407–430.
Fiedler, K. (1996). Explaining and simulating judgment biases as an aggregation phenomenon in probabilistic, multiple-cue environments. *Psychological Review, 103,* 193–214.
Fodor, J. A. (1995). The folly of simulation. In P. Baumgartner & S. Payr (Eds.), *Speaking minds: Interviews with twenty eminent cognitive scientists* (pp. 85–100). Princeton, NJ: Princeton University Press.
Gigerenzer, G. (1981). *Messung und Modellbildung in der Psychologie* [Measurement and modelling in psychology]. Munich: Ernst Reinhardt (UTB).
Gigerenzer, G. (1992). Discovery in cognitive psychology: New tools inspire new theories. *Science in Context, 5,* 329–350.
Gigerenzer, G. (1993). The superego, the ego, and the id in statistical reasoning. In G. Keren & C. Lewis (Eds.), *A handbook for data analysis in the behavioral sciences: Methodological issues* (pp. 311–339). Hillsdale, NJ: Erlbaum.
Gigerenzer, G. (1994). Why the distinction between single-event probabilities and frequencies is relevant for psychology (and vice versa). In G. Wright & P. Ayton (Eds.), *Subjective probability.* Chichester: Wiley.
Gigerenzer, G. (1996). On narrow norms and vague heuristics: A rebuttal to Kahneman and Tversky (1996). *Psychological Review, 103,* 592–596.
Gigerenzer, G., & Murray, D. J. (1987). *Cognition as intuitive statistics.* Hillsdale, NJ: Erlbaum.
Gigerenzer. G., & Regier, T. (1996). How do we tell an association from a rule? Comment on Sloman (1996). *Psychological Bulletin, 119,* 23–26.
Gigerenzer, G., Swijtink, Z., Porter, T., Daston, L. J., Beatty, J., & Krüger, L. (1989). *The empire of chance: How probability changed science and everyday life.* Cambridge, England: Cambridge University Press.
Kahneman, D., & Tversky, A. (1996). On the reality of cognitive illusions. *Psychological Review, 103,* 582–591.
Massaro, D. W. (1988). Some criticisms of connectionist models of human performance. *Journal of Memory and Language, 27,* 213–234.
Newell, A. (1973). You can't play 20 questions with nature and win: Protective comments on the papers of this symposium. In W. G. Chase (Ed.), *Visual information processing* (pp. 283–308). New York: Academic Press.
Oaksford, M., & Chater, N. (1994). A rational analysis of the selection task as optimal data selection. *Psychological Review, 101,* 608–631.
Regier, T. (1996). *The human semantic potential: Spatial language and constrained connectionism.* Cambridge, MA: MIT Press.
Sahlin, N.-E. (1991). Baconian inductivism in research on human decision-making. *Theory & Psychology, 1,* 431–450.
Shanteau, J. (1989). Cognitive heuristics and biases in behavioral auditing: Review, comments and observations. *Accounting Organizations and Society, 14,* 165–177.
Shepard, R. N. (1962). The analysis of proximities. Part 2. *Psychometrika, 27,* 219–246.
Shepard, R. N. (1974). Representation of structure in similarity data: Problems and prospects. *Psychometrika, 39,* 373–421.
Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin, 119,* 3–22.
Smith, E. E., & Osherson, D. N. (1989). Similarity and decision making. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 60–75). Cambridge, England: Cambridge University Press.
Tooby, J., & Cosmides, L. (1992). The psychological foundations of culture. In J. H. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture.* New York: Oxford University Press.
Tversky, A. (1977). Features of similarity. *Psychological Review, 84,* 327–352.
Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185,* 1124–1131.
Wallach, L., & Wallach, M. A. (1994). Gergen versus the mainstream: Are hypotheses in social psychology subject to empirical test? *Journal of Personality and Social Psychology, 67,* 233–242.
Wallach, M. A., & Wallach, L. (1998). When experiments serve little purpose: Misguided research in mainstream psychology. *Theory & Psychology, 8,* 183–194.
Wallsten, T. S. (1983). The theoretical status of judgmental heuristics. In R. W. Scholz (Ed.), *Decision making under uncertainty* (pp. 21–39). Amsterdam: Elsevier.
Wanke, M., Schwarz, N., & Bless, H. (1995). The availability heuristic revisited: Experienced ease of retrieval in mundane frequency estimates. *Acta Psychologica, 89,* 83–90.

Acknowledgements. I am grateful to Valerie Chase for comments on an earlier version of this paper.

Gerd Gigerenzer has been Director at the Max Planck Institute for Human Development in Berlin since 1997. He was previously Director at the Max Planck Institute for Psychological Research, Munich, and Professor of Psychology at the University of Chicago. His work centers on bounded rationality, social rationality, and judgment and decision making under uncertainty. He is the 1992 winner of the AAAS prize for Behavioral Science Research, and the (co-)author of several books, including *Cognition as Intuitive Statistics* (Erlbaum, 1987) and *The Empire of Chance: How Probability Changed Science and Everyday Life* (Cambridge University Press, 1989). He is currently writing a book on bounded rationality entitled *Simple Heuristics That Make Us Smart* (with Peter M. Todd and the ABC Group: Oxford University Press). Last, but not least, he is coeditor of *Theory & Psychology.* Address: Center for Adaptive Behavior and Cognition, Max Planck Institute for Human Development, Lentzeallee 94, 14195 Berlin, Germany. [e-mail: gigerenzer@mpib-berlin.mpg.de]