

Representative Design

Kenneth R. Hammond
Professor Emeritus, Department of Psychology
University of Colorado

What, exactly is "representative design"? Why did Brunswik introduce this idea in the first place? And what difference does it make? I will try to answer these questions briefly in this Web-Essay. And because this topic has been discussed many times (see for example, Brehmer & Joyce, 1988; Cooksey, 1996; Hammond, 1954, 1966, 1996; not to mention Brunswik, 1943, 1955, 1956), I will try not to repeat exactly what has already been said so many times, but offer a different approach.

Fundamental to the principle is the rule that one may generalize the results of observations only to those circumstances or objects that have been sampled.

What is representative design? Simply put, it is a principle for the design of experiments that adheres to the logic of inductive inference as it is employed by philosophers, psychologists, and statisticians. This principle is based on the sampling theory that gives modern inductive inference its operational basis. Fundamental to the principle is the rule that one may generalize the results of observations only to those circumstances or objects that have been sampled. And that is the idea that Brunswik applied to the design of experiments. Of course, that idea had been applied for some time by psychologists and statisticians to the *subjects*, that is, participants, in the experiments (or surveys). This principle was not only accepted but became mandatory science.

This was the first time a conceptual effort of this kind had been made, and it is that kind of hard intellectual work that sets the idea of representative design off from the mindless use of the term "real life."

Brunswik's contribution was novel, however; he pointed out that psychologists were operating under a "double standard." Sampling theory was being ignored entirely with regard to the *objects*, or more generally, the "stimulus," or "input," or environmental *conditions* of the experiment. Why, he wanted to know, is the logic we demand for generalization over the subject side ignored when we consider the input or environment side? Why is it that psychologists scrutinize subject sampling procedures carefully but cheerfully generalize their results--without any logical defense--to conditions outside those used in the laboratory. He made his argument in no uncertain terms, and gave detail after detail, particularly in his 1956 book, where he offered a schema for classifying "the variables entering objective psychological research . . . with respect to regions relative to an organism" (p. 4.) This was the first time a conceptual effort of this kind had been made, and it is that kind of hard intellectual work that sets the idea of representative design off from the mindless use of the term "real life." It is important to note that this "classification of the variables entering objective psychological research . . . with respect to regions relative to an organism" constitutes a *theory of the environment*. That makes Brunswik unique among psychologists for the "classification" is an explicit statement naming variables in a formal way. And that classification provides the basis for examining the variables that any experiment includes, and thus makes explicit the permissible generalization. I will say more about this in my Web-Piece on the misuse of the term "ecological validity," but those readers who have access to the 1956 book can turn to the appropriate pages. (Note: I recommend the first 38 pages of this book. These will not be easy reading, but they are illuminating; they show the history of design in psychology, and also show the operational features of representative design.)

Although Brunswik's thesis--put forward with vigor in the *Psychological Review* article in 1943--contained no *ad hominem* remarks, it drew emotional responses, seldom published, from the most well-known psychologists of the day. (I have received plenty of these never published scornful remarks myself.) Only on one occasion, at the Unity of Science meeting in Berkeley in 1953, were Brunswik's ideas challenged openly by prominent psychologists, largely at his invitation. It will pay those interested in Brunswik's ideas to read these challenges as well as his paper and his rebuttal to his critics; all are in the *Psychological Review* article (1955); there are lucid expressions of his defense of probabilistic functionalism and representative design as well.

Experimental psychologists have made two main objections to Brunswik's argument.

Experimental psychologists have made two main objections to Brunswik's argument. The first stems from its innovative character; you won't find a mention of the demand for a consistent application of a theory of induction for the results of experiments before Brunswik brought it up in a mainstream article in 1943. This leads to the objection that we shouldn't bother with this matter because it can't be done, or at least it can't be done easily. How can one possibly sample situations? (This objection persists to this day: See Dawes, 1998, in *The Handbook of Social Psychology*.) The second objection is that we need not bother with this matter, anyway, even if we could do it, because we don't *want* to do it. That is, we psychologists aren't interested in generalizing our results from experiments to conditions outside the experiment, nor should we be. We just want to find out what happens in the conditions we deem necessary to test our hypotheses; get the experiment right; that's all that matters.

A pursuit of the first objection was once challenging but is becoming less so as simulation of task circumstances in the conditions of interest becomes more and more available. For example, studying weather forecasters in their work environment becomes easier and easier as it becomes easier to replay the tapes of their radar representations of weather; the same is true of displays of medical data, military data, aviation instrument data, industrial systems data, etc. Moreover, we can now create a variety of *new*--"what if"--situations of interest to us on tapes and display them. And as more and more information is displayed to us via tapes and computer screens, representation of conditions of interest by video techniques will become the rule rather than the exception. As this happens the debate over representative design will fall by the wayside; almost everyone will be able to carry out their work by displaying conditions representative of the conditions toward which the generalization is intended. Students of the next generation who are able to free themselves from their professor's "methodological ideology" will wonder what the fuss was about.

The second objection is the most difficult to cope with. I think we are at an impasse here. But if we are indeed at an impasse we should recognize that, and editors and authors should make clear their positions on the matter. What should *not* be continued is the current practice of nonsensical claims, namely, that a given experiment does or does not reflect "real life" or has a "lot of ecological validity." But let me try once more to make the case for defensible generalizations by taking a simple example that (I apologize) has been used many times before.

Consider the topic of person perception. Let us assume we want to assess the accuracy of our judgments of traits of other people. Our study will require that we gather a group of persons (subjects, or participants as they are now called) and ask them to make judgments about the traits (friendliness, etc.) of a number of other people. Brunswik called the persons to be judged "person-objects" to distinguish them from the "subjects" in the experiment. In the past--and the recent

past--psychologists would gather anywhere from 50 to several hundred persons to be subjects in the experiment, to make the judgments about the person objects. They used so many subjects because they wanted a secure generalization over the subjects.

But it was not uncommon--nor is it today--to find psychologists ignoring the question of the legitimacy of generalizing over the person objects. We know that they are ignoring that question because they all too often used only *one person* as a person object. Despite this flagrant disregard of sampling theory, they would report their findings about the ability of people (in general) to make accurate judgments about other *persons*, also in general, despite their sample of *one*. (One of my early papers (Hammond, 1954) demonstrated how a very reputable psychologist carried out a prominent study using a large subject sample but a person object sample of *two*. in order to discover that effect of the examiner [person object] on a subject's responses.) This error, this "double standard," was so widespread in the 1940s, 1950s, and 1960s that I have been tempted to make a collection of these studies in a book, just to show what can happen to a science that innocently gets off on a wrong track. And it still happens. Several times a year the *Journal of Personality and Social Psychology* will publish a study of this sort. I can only conclude that the editor, reviewers, and author(s) are ignorant of the topic.

In short, you have to be specific about the variables that a *theory* tells you are the important ones, namely, the ones that if ignored would produce a critically different result than if not ignored. These were laid out by Brunswik in his "classification of variables relative to an organism."

But that was an easy case in which to see the point. But how about the situation where, say, learning theorists are running rats in mazes? How in the world is it possible to *sample* rat mazes? Well, first of all, as in any good sampling procedure, you have to be specific about the criteria of interest. In opinion polling studies these criteria are related to the content of the study and usually include such demographics as age, ethnic groups, gender and the like. In short, you have to be specific about the variables that a *theory* tells you are the important ones, namely, the ones that if ignored would produce a critically different result than if not ignored. These were laid out by Brunswik in his "classification of variables relative to an organism."

That brings us to a theory of mazes, which is to say, a theory of rat environments. That is precisely what the learning theorists of the day did *not* present (a theory of mazes would have seemed bizarre to them) but it is precisely

what Brunswik *did* present to them in 1939, and it did seem bizarre. But because his theory of the environment included the idea of uncertainty, that basic idea became the crucial variable in his theory of the maze. So in 1939 he offered a specific example of what he meant by "uncertainty in the environment" in his one and only study of rat learning ("Probability as a Determiner of Rat Behavior"). By demonstrating the effect of uncertainty, and discovering a threshold for uncertainty affecting learning in the environment, he became one of the first psychologists ever to do so. Ever alert to the politics of scientific acceptability, Brunswik must have realized that if he was to gain the attention of the learning theorists--the psychologists who mattered in those days--he would have to demonstrate his argument about probability on their terms, that is, in the rat laboratory. He made his point in the first paragraph: "In the natural environment of a living being, cues, means or pathways to a goal are usually neither absolutely *reliable* (italics added) nor absolutely wrong. In most cases there is, objectively speaking, no perfect certainty that this or that will, or will not, lead to a certain end, but only a higher or lesser degree of probability (cf. Tolman & Brunswik)." In spite of this commonplace observation, Brunswik pointed out, the mazes of the day were arranged so that "usually the connection between means and end is made by the experimenter to be what Hume or John Stuart Mill would call indissoluble or inseparable, one of the alternative behaviors always being rewarded and the other never" (p.175). (This sentence is typical of Brunswik's style; he frequently reached back into the history of philosophy and/or psychology to call the reader's attention to the significance of a point. This practice, I have observed, irritates experimental psychologists considerably, for reasons I leave to the reader's conjecture.) Brunswik's suggestion about the ubiquity of environmental uncertainty will seem commonplace to today's psychologist (or anybody else), but the learning theorists in 1939 did not find it persuasive. Not only was it a new idea but when presented by Brunswik it was firmly rejected by the major theorists of the day (Hull, Spence, Lewin, Kofka, Koehler, and others; see the 1941 symposium that included Hull, Lewin and Brunswik in which Brunswik makes uncertainty and probability the centerpiece of his paper, although ruled out of psychology by Hull and Lewin.)

The results of his experiment were clear. Equally clear was the response of the learning theory psychologists: Ignore this suggestion and its demonstration, or if faced with it as Clark Hull was when confronted with Brunswik in the 1941 symposium, deny its relevance by denying the place of "probability laws" in psychology, and say that anyone who urges that we make use of them must have given up, "as Brunswik seems to have done." Some twenty-five years later, however, it was Clark Hull and his form of learning theory that was "given up." It wasn't killed by a competing theory; rather, its obvious irrelevance to the learning of organisms in their natural environments became all too apparent, and interest in it just faded away.

(Note: I italicized the word "reliable" in the above quoted paragraph to call to the reader's attention the point that in 1939 Brunswik applies the word

"reliable" to a cue in the colloquial sense. He does not use the word "valid" because, I suspect, he had not yet heard of it. He did this experiment in 1938 during his first year as a member of the faculty at Berkeley, where he would have to learn about psychometrics. Statistics were not a part of psychology in Europe at that time; the first reported use of the correlation coefficient being in the mid-thirties. Thus it is clear that in 1939 he had not yet developed the concept of "ecological *validity* ." (I will say more about this development in my next piece which concerns the current status of the concept of ecological validity.)

One can understand resistance to new ideas such as "probability" and "probability laws" in the 1930s , 1940s, and 1950s. We no longer hear much about "laws" in psychology today, but we hear a great deal about "probability" and "uncertainty," and certainly Brunswik deserves credit for his groundbreaking theory and research in that connection, but seldom, if ever gets it.

I believe that a good part of the resistance to accepting Brunswik's argument by experimental psychologists lies in their flawed education; they typically don't learn about representative design until they are mid-career (if then) and by that time have published several articles that fail the representativeness test.

I believe that a good part of the resistance to accepting Brunswik's argument by experimental psychologists lies in their flawed education; they typically don't learn about representative design until they are mid-career (if then) and by that time have published several articles that fail the representativeness test. At that point it's hard to say: "Oh, I see, I was mistaken in those four studies in making the claims I did. I take it all back." No one can bring him/herself to do that. But since the principles of representative design have never been explained by the textbook writers. psychologists are generally introduced to it in a happenstance way. Secondly, many arguments offered *for* representativeness are highly informal, and worse, based on a misunderstanding. That misunderstanding is brought about by those who abuse and debase the concept of "ecological validity," to which I will turn in my third Web-Piece. It is now commonplace to find an author claiming that his or her work is "ecologically valid," or that someone else's isn't, and to substantiate that claim by arguing that his or her experimental conditions somehow resemble those of the "real world"--an absurd concept--while the others do not. This has resulted in such semantic confusion, to be overly polite, that we find Robyn Dawes complaining about "*too much* ecological validity" while others complain about too little. Somehow, about 25 years ago someone, I'm not sure who, casually used the term "ecological validity" to mean that his or her experimental results would--somehow--generalize beyond the confines of the laboratory. And the usage became popular and stuck; its

original meaning--the relation between a cue and a distal variable--had been changed to confuse it with another central concept, namely, representative design.

There is great irony--and sadness--in those events, in the misidentification of "ecological validity" with the representative design of experiments. The irony lies in the general disparagement of Brunswik's concept of representative design (when it is recognized) and the now widespread demand for ecological validity, which is usually intended to mean "representativeness" of conditions outside the laboratory. The irony is stunning, because the effort to achieve "ecological validity" is, in fact, an effort to achieve some sort of representative design, and thus presents an empirical vindication of Brunswik's advocacy of representative design, yet the effort is based on a corruption of both the concept of ecological validity and representative design, a corruption that comes about largely through ignorance. Thus, the irony is compounded: The corrupted concept of representative design is slowly but surely earning a place in the methodology of modern psychology, but doing so under the guise of the corrupted concept of ecological validity. No doubt the term ecological validity --in its corrupted form, of course--will soon appear in textbooks that explain the importance--or unimportance--of linking one's research to the "real world," whatever that is. When such confusion is rampant in the methodological conventions of psychology, it is small wonder that psychology is having difficulty in becoming a cumulative science.

References

- Brehmer, B., & Joyce, C. R. B. (Eds.). (1988). *Human judgment: The SJT view*. Amsterdam: Elsevier.
- Brunswik, E. (1939). Probability as a determiner of rat behavior. *Journal of Experimental Psychology*, 25, 175-197.
- Brunswik, E. (1943). Organismic achievement and environmental probability. *Psychological Review*, 50, 255-272.
- Brunswik, E. (1955). Representative design and probabilistic theory in a functional psychology. *Psychological Review*, 62, 193-217.
- Brunswik, E. (1956). *Perception and the representative design of psychological experiments* (2nd ed.). Berkeley, CA: University of California Press.
- Cooksey, R. W. (1996). *Judgment analysis: Theory, methods, and applications*. San Diego: Academic Press.
- Dawes, R. M. (1998). Behavioral decision making. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (Vol. 1, pp. 497-548). Boston: McGraw-Hill; Distributed exclusively by Oxford University Press.
- Hammond, K. R. (1954). Representative vs. systematic design in clinical psychology. *Psychological Bulletin*, 51(2), 150-159.

Hammond, K. R. (Ed.). (1966). *The psychology of Egon Brunswik*. New York: Holt, Rinehart, and Winston.

Hammond, K. R. (1996). *Human judgment and social policy: Irreducible uncertainty, inevitable error, unavoidable injustice*. New York: Oxford University Press.