

PSYCHOLOGY'S SCIENTIFIC REVOLUTION:  
IS IT IN DANGER?

by  
Kenneth R. Hammond<sup>1, 2</sup>

UNIVERSITY OF COLORADO  
INSTITUTE OF BEHAVIORAL SCIENCE

Center for Research on Judgment and Policy

Report No. 211

May 1978

## Abstract

Brunswik's revolutionary approach to psychology has steadily increased its influence over the past three decades. Unfortunately, as the use of this approach increases, its principal concepts, representative design and ecological validity, are being confused, corrupted and reduced to cliches, thus endangering further progress. In an effort to restore and preserve their meaning, these concepts are described and examples of their use and misuse are provided. To illustrate further the potential significance of the loss of these concepts, the widely-accepted phenomenon of "illusory correlation" is critically examined to show that because it was studied without the benefit of the above mentioned concepts it remains without theoretical or empirical foundation.

### Psychology's Scientific Revolution: Is It in Danger?

A scientific revolution is underway in psychology. It began when a prominent scientist (Brunswik) argued that the conventional approach within his discipline is the reverse of what it ought to be and provided cogent theoretical and empirical support for the argument. More and more research workers are finding the conventional approach inadequate, are calling for the same reversal in approach, and employing the new theory and methodology in their work. When such events occur the discipline is, in the well-known language of Thomas Kuhn (1962), giving up one paradigm for another.

Psychology's scientific revolution began with Brunswik's argument for a reversal in psychology's approach to its subject matter. He (a) urged the abandonment of the search for nomothetic, deterministic laws of behavior in favor of idiographic-statistical descriptions (now "models") of the behavior of individuals, (b) advocated the replacement of the systematic design of experiments in favor of the design of experiments that are representative of the organism's ecology, or habitat, and (c) suggested that the geographer's descriptive efforts rather than the physicist's law-seeking endeavor should provide the proper model for psychologists (1943, 1952, 1956). These points were buttressed by a detailed methodological argument, empirical studies and a sophisticated analysis of trends in the history of psychology which

pointed to the eventual acceptance of these revolutionary ideas. (See Hammond, 1966 for an overview.)

The reversal in approach urged by Brunswik is now being advocated by an increasing number of psychologists. From the time of Koch's (1959) review of psychology in which he asserted that there is a "stubborn refusal of psychological findings to yield to empirical generalization," continuing through the present to Cronbach's gloomy pronouncement that "generalizations decay" (1975, p. 122), there has been a growing recognition that the present paradigm is failing; that change is needed. Leaders in the fields of experimental psychology (e.g., Jenkins, 1974), developmental psychology (e.g., Bronfenbrenner, 1977), and educational psychology (e.g., Cronbach, 1975), and social psychology in particular (e.g., Gergen, 1973, 1976; McGuire, 1973; Smith, 1976; see also Elms, 1975) have noted that the traditional approach leaves us with results that are restricted to the laboratory, and therefore of dubious value. Increasingly, the general demand is for some methodological change (often no more than brave, if vague, calls for more "field research") that will permit the achievement of results that will carry stable meaning for behavior outside the laboratory (Reference Note 1).

As the pace of change increases, however, the principal concepts of psychology's scientific revolution are becoming confused, emptied of meaning, reinvented, and their origins lost. In the hope of preventing further negative developments of this kind, I shall (a) indicate briefly the nature of certain key Brunswikian concepts, (b) show that they are in danger of losing their meaning as their use increases, (c) call attention to the original meaning of these concepts in an effort to

restore their theoretical coherence, and (d) demonstrate how false, yet widely accepted, conclusions have been drawn from research that fails to maintain the integrity of these concepts. It must be emphasized that it is not merely the priority of usage of terms that is at stake; rather, it is my purpose to serve the development of psychology as a cumulative science, and to spare its students from having to cope with the idiosyncratic proliferation of the meanings of concepts taken from what was originally a theoretically coherent context. Since it is nearly twenty-five years since Brunswik's death, however, his academic background is briefly described first.

#### Egon Brunswik

Brunswik (1903-1955) began his work during the 30's in Austria and continued it at the University of California at Berkeley from 1937 to 1955. Although Brunswik earned high esteem as a scholar for his profound analyses of the history, method and theory of psychology, as well as his research, his arguments for change were rejected during his lifetime (see Hammond, 1968; Tolman, 1956, for biographical information; for contrasting views with Hull and Lewin, see Brunswik, 1943; Hull, 1943; Lewin, 1943).

To what extent this rejection accounted for his suicide in 1955 is uncertain; what is certain is that the acceptance of his ideas has increased steadily since his death; there have been over 175 references to his work in the last five years. What is equally certain is that he called for a reversal of fundamental practices. As noted by Gibson (1957) in his review of Brunswik's book (1956): "He [Brunswik] asks us. . . to revamp our fundamental thinking. . . . It is an onerous

demand. . . . His work is an object lesson in theoretical integrity" (p. 35). Only by maintaining that "theoretical integrity" will psychology be able to maintain its momentum toward the attainment of laboratory results that carry meaning for situations outside the laboratory.

### Three Revolutionary Concepts in Danger

The current treatment of three concepts introduced by Brunswik (representative design, ecological validity, and intra-ecological correlation) are discussed below. In what follows we show how the current treatment of these concepts produce the confusion, loss of origin and meaning, and reinvention mentioned above.

#### Representative Design

The concept of representativeness is fundamental to generalization. Just as the subjects in an experiment must represent those not included in the experiment if generalization over subjects is to be achieved, so also must the conditions of an experiment represent those conditions outside the laboratory over which generalization is to be achieved.

Systematic arrangement of conditions in the experiment that do not represent the nonsystematic arrangement of conditions outside of it prevents both logical and empirical generalization of results. Moreover, if experiments are to produce results that will generalize to circumstances outside the laboratory, they must not merely include substantive material that is representative of the outside situation, but the formal, that is, structural, aspects of the situation outside the laboratory as well. As Brunswik put it: "Generalizability of results concerning. . . the variables involved [in the experiment] must remain limited unless the range, but better also the

distribution. . . of each variable, has been made representative of a carefully defined set of conditions" (1956, p. 53). Brunswik's admonition regarding the representativeness of the formal aspects of the conditions of experiments also includes the (ecological) inter-correlation among the independent variables in the experiment, thus challenging the typical factorial design in which variables are set in orthogonal relation to one another.<sup>3</sup>

One way to achieve representativeness of task, or environmental, conditions is to sample conditions, just as one achieves representativeness of the subject population by sampling subjects. In this way the range, distribution and inter-correlation among environmental variables will appear in the laboratory sample, and, therefore the laboratory conditions will be representative of the conditions toward which generalization is intended--within the limits of sampling errors. The size of such sampling errors can be estimated and controlled by the size of the sample, precisely as in subject sampling.

This argument for the need for representativeness of task conditions has slowly but surely gained acknowledgement since Brunswik introduced it in the 40's (see, e.g., Brunswik, 1943), particularly in the area of person-perception. For when task conditions include persons, the argument is straightforward and its implementation is not difficult. It is easy to see that if one intends to generalize over persons, whether they are subject-persons or object-persons, sampling is required. If one wants to ascertain whether short people are perceived to be less (or more) aggressive than tall people one must expose to the subject sample an adequate sample of short and tall person-objects since short and tall persons will vary in many other dimensions, e.g.,

weight, sex, etc. And, indeed, object-person sampling has increased substantially in research on social perception and social judgment in the last two decades. From a time during the 50's when virtually all studies in social perception and social judgment ignored object sampling (see Crow, 1957; Hammond, 1948, 1954 for examples in this period) and thus made their results useless, research in social judgment and social perception has progressed to the point where nearly all researchers now employ reasonably sized numbers, if not true samples, of persons-objects, or target persons. This aspect of the revolution is almost secure.

The use of systematic design has not altogether been given up, however. Studies which employ intensive subject-person sampling but no object-person sampling are still published in the Journal of Personality and Social Psychology as well as other prominent journals. In 1977, for example, Nisbett and Wilson (1977) had 118 subjects rate a single object-person, and Selby, Calhoun, and Brock (1977) had 47 subjects rate one object-person. The most recent issue to arrive (April, 1978) contains a study of lying that includes a single "liar" (Kraut, 1978). A sobering fact is that when Thorndike (1920) discovered the "halo effect," over a half-century ago, he got it right (there were 8 subjects and 137 object-persons in the study he described) as did the agronomist, Henry Wallace, when he studied the corn judge's ratings of ears of corn in 1923. The price psychology pays for clinging to the conventional systematic research paradigm is evidenced in a review of research on the effect of the sex of the experimenter in psychological research (Rumenik, Capasso, & Hendrick, 1977). They found that only 39 of 63 studies used as many as two members of each sex as experimenters (p. 874) and



few used more than that. (Only 8 of the 63 studies used as many as 10 in each sex group!) Many of the studies that failed to employ object (experimenter) sampling were conducted in the 1970's. Such wasted effort should not be surprising, however; students are instructed in a fashion that perpetuates these errors (see, for example, Aronson & Carlsmith, 1968) by researchers whose own work provides examples of the same mistakes (e.g., Aronson, Willerman, & Floyd, 1966).

The choice between representative and systematic design of experiments has surfaced episodically since the beginning of scientific research in psychology. As Gillis and Schneider (1966) point out, Wundt recognized it in 1896 and chose systematic design; MacDougall recognized it in 1922 and chose representative design; Wundt was followed and MacDougall was not. When Brunswik recognized the necessity for representative design, he laid out in great detail the profound implications of that choice for theory and method, and for the nature of psychology as a scientific discipline (1952, 1956). He went so far as to suggest that choosing representativeness in experimental design would mean giving up the law-seeking physicist as the role model for psychologists and choosing the description-seeking geographer instead.

The revolutionary implications of this idea are still imperfectly understood because they have not yet been fully explored. But it is a revolutionary idea that is gaining acceptance (Cronbach, 1975). Its acceptance will be unduly slowed, however, if the specific meanings of revolutionary concepts are eroded, for psychologists and their students will then be required to relive their history in a

morass of confused and indistinct concepts and amidst studies that produce results of doubtful utility--a situation not far from that which exists today. The erosion of meaning from the concept of ecological validity and its confusion with the concept of representative design (and its aim of generalization) provide clear examples of these dangers.

### Ecological Validity

Brunswik introduced the term ecological validity to indicate the degree of correlation between a proximal (e.g., retinal) cue and the distal (e.g., object) variable to which it is related (see Brunswik, 1956, pp. 48-52, on the "Ecological Validity of Potential Cues and Their Utilization in Perception"). Thus, in a perceptual task, ecological validity refers to the objectively measured correlation between, say, vertical position and size of an object (larger objects tend to be higher up in the visual field) over a series of situations. Or, more broadly, one may compare the ecological validity of the cue "height of forehead" with the cue "vocabulary level" as indicators of a person-object's intelligence. (See, for example, the section entitled "Classification of cues in terms of ecological validity, representative design," Brunswik, 1957, or "Ecological validity of potential cues and their utilization in perception," pp. 48-50, 1956.) In short, ecological validity refers to the potential utility of various cues for organisms in their ecology (or natural habitat). Of course, the difference between the ecological validity of a cue and its actual use by an organism provides important information about the effective use of information by that organism (see Fig. 1).

### Erosion of Meaning

Prior to 1974, the concepts of representative design and ecological validity were kept distinct, and psychologists used them as they were intended to be used by their author. Among those who employed these terms as Brunswik defined them are Bruner, Goodnow, and Austin, 1956; Dudycha and Naylor, 1966; Einhorn, 1972; J. Gibson, 1957; Gillis, 1975; Goldberg, 1970; Hammond, 1955; Hammond, Rohrbaugh, Mumpower and Adelman, 1977; Hammond, Stewart, Brehmer, and Steinmann, 1975; Heider, 1958; Hochberg, 1966; Jarvik, 1966; Keeley and Doherty, 1972; Leeper, 1966; Lewin, 1943; Lindell, 1976; Loevinger, 1966; Murell, 1977; Osgood, 1957; Postman and Tolman, 1959; Rappoport and Summers, 1973; Slovic and Lichtenstein, 1971, Smedslund, 1955; Steinmann and Doherty, 1972; Stewart, 1976; and several authors of chapters in the Handbook of Social Psychology (1968, see, e.g., Tajfel, pp. 315-394) and others. In short, these terms have had established meaning for over three decades (1947-1977). Their meanings are no longer unique to Brunswik, and for the many research workers who have used these terms in a precise way their meanings are not arbitrary. To assign new meanings without reference to previous use not only introduces confusion, it is bad science. Once a term loses its meaning, it is virtually impossible to recover it (cf. Hochberg, 1956).

Unfortunately, however, we find that since 1974 the term ecological validity has been used in a number of very different ways by different authors who do not recognize either Brunswik's use of the term or anyone else's use of it. Thus, Jenkins (1974) talks about the "ecologically valid problems of everyday life." Bronfenbrenner (1977) refers to the

ecological validity of experiments, as do Berman and Kenny (1976), as well as Graham (1977) who, after mistakenly attributing the concept to Orne (1970) (who mistakenly refers to "Egon Brunswik's concept of the ecological validity of research" [p. 259]) defines this term as "the extent to which the setting in which research takes place is capable of producing results that are valid." Neisser (1976) on the other hand, refers to the ecological validity of theories. And a number of authors (Christensen, 1977; Eaton & Clore, 1975; Frodi, 1974; Greenwald, 1976; Orne, 1970; Parke, 1976; Silverstein & Strang, 1976) refer to the ecological validity of results.

Erosion of meaning from the concept of ecological validity is perhaps best illustrated by Neisser's use of it (1976, p. 48). Neisser acknowledges that the term "ecological validity was coined by Brunswik," but adds that Brunswik's use of the term "was slightly different from the one that is popular today." Unfortunately, Neisser offers neither Brunswik's definition of the term, the one that is "popular today," nor his own, and thus empties the concept of meaning.

Since 1974, the term ecological validity has come to be used by some authors to refer to the degree to which results obtained in the psychological laboratory "generalize" to circumstances outside the laboratory. Jenkins, for example, in his 1974 presidential address to members of the Division of Experimental Psychology admonishes them that "it is true. . . that a whole theory of an experiment can be elaborated without contributing in an important way to the science because the situation is artificial and non-representative. . . . In short, contextualism stresses relating one's laboratory problems to

the ecologically valid problems of everyday life" (p. 794). Readers should not let the puzzling phrase "ecologically valid problems of everyday life" distract them from the main point: Jenkins confuses the concept of ecological validity with generalization and the representative design of experiments.

Bronfenbrenner (1977), in his presidential address to the members of the Division of Personality and Social Psychology in 1974, also confuses the concept of ecological validity with generalization and representative design. He begins his critique of current and past psychological research by observing that "the emphasis on rigor has led to experiments that are elegantly designed but often limited in scope. . . . Many of these experiments involve situations that are unfamiliar, artificial, and short-lived and that call for unusual behaviors that are difficult to generalize (italics ours) to other settings" (p. 513). Having expressed his dissatisfaction with the lack of representativeness of past and current research design which makes it difficult to generalize laboratory findings to non-laboratory situations, Bronfenbrenner turns to the concept of ecological validity (p. 515) and states that "although this term has, as yet, no accepted definition" (thus joining with Neisser and Jenkins in ignoring three decades of empirical research and a substantial body of psychological theory) he then proceeds to change the established definition of ecological validity by saying that "one can infer from discussions of the topic a common underlying conception: An investigation is ecologically valid if it is carried out in a naturalistic setting and involves objects and activities from everyday life." Finding his own new definition not only "too simplistic" and "scientifically unsound," "as it is currently used" (no reference), he also finds it to have

"no logical relation to the classical definition of validity--namely, the extent to which a research procedure measures what it is supposed to measure." This statement, even in its idiosyncratic form, is simply false. In the articles written previous to 1974, the concept of ecological validity was consistently used within the classical definition of validity, as it should be. Indeed, in his Table 2, on p. 30, Brunswik (1956) not only defines ecological validity in test measurement terms, but defines the ecological reliability of cues in test measurement terms also, thus preserving the kind of theoretical coherence any science requires if it is to be cumulative.

Ignoring all past work on the representative design of experiments and the ecological validity of cues, and finding his own redefinition wanting, Bronfenbrenner then offers a second new definition of ecological validity, namely; "Ecological validity refers to the extent to which the environment experienced by the subjects in a scientific investigation has the properties it is supposed or assumed to have by the investigator" (p. 416). Needless to say, all connection with previous usages is now hopelessly lost.

The consequences of steady erosion of meaning from this concept are now beginning to appear; equally prestigious psychologists are now protesting against "indiscriminate use" of the term and implying that its "indiscriminate use" has already destroyed its value. Bandura, for example, (1978) has protested against "the slighting of experimentation by recourse to the . . . ready invocation of ecological validity. This notion," he correctly observes, "has lost much of its identity from its earlier parentage, (and) is in danger of being transformed into a cliché through indiscriminate use."

In short, despite three decades of consistent and growing use, the established meanings of these valuable concepts of representative design and ecological validity are being eroded, confused, changed arbitrarily. Indeed, these concepts may have already been corrupted beyond retrieval, perhaps to the point where their abandonment could become necessary because they will have become reduced to cliches, and their scientific value thereby lost, circumstances Bandura sees as imminent. If that should happen, psychologists will be deprived of concepts already proven to be so useful that they will have to be reinvented.

From a dictionary or common language point of view, Neisser, Jenkins and Bronfenbrenner, or any one else is free to use the terms ecological validity and representative design in whatever fashion ordinary usage will permit. But scientists do not use the common language in the common way; they assign special definitions to specific terms. And when that occurs, and when those special definitions acquire stable and significant meanings to many workers in the discipline over a period of time, then indifference to established usage together with arbitrary redefinition become obstacles to progress. The discipline that permits such obstacles cannot be a cumulative science--a matter about which psychology has reason to be tender.

#### Intra-ecological Correlations

The term intra-ecological correlation is meant to apply to the co-variation among variables in the social, biological and physical ecology of organisms. This concept is basic to the Brunswikian revolution (see Section VI, 1956). Historically, intra-ecological

correlations were, of course, the first aspects of real world, ecological conditions to be brought under control (that is, reduced to zero) in the attempt to achieve precision, best illustrated in the traditional psychophysical research paradigm. Brunswik shows (pp. 12-24) how such "controls" over ecological correlations were gradually relaxed as psychologists attempted to extend the generality of their findings. Variations in the length of the feathers and arrows of the Muller-Lyer illusion, for example, demonstrated that context can markedly affect perceptions and judgments of even the simplest stimuli. When it became clear that contextual variables can markedly affect psychological processes, methodological imperatives were reversed and social psychologists, in particular, began to investigate the question of whether contextual variables do affect psychological processes. For if they do, then results obtained in the laboratory under conditions of control that reduce such correlations to zero will not generalize to those conditions outside the laboratory in which intra-ecological correlations are not zero.

The concept of intra-ecological correlation is, therefore, fundamental to the logical step of generalization from the laboratory to the real world, for such correlations must be present in the laboratory task in the same degree as in the real world if generalizations are to be made from one to the other. (See Hammond, 1966, pp. 29-30 for further discussion of the importance of intra-ecological correlations in Brunswik's methodology; see also Hammond, Stewart, Brehmer, & Steinmann, 1975, for a recent discussion). And although this Brunswikian concept is gradually becoming accepted, as the research



to be cited below will indicate, failure to employ this concept leads to unfortunate results (as in the case of object sampling described above). The final section of this article presents an example of how a series of studies have led to the wide acceptance of conclusions that are not supportable because the authors confuse representativeness of design, ecological validity, and intra-ecological correlation.

### "Illusory Correlation"

In 1967 and 1969 three articles appeared (Chapman, 1967 and Chapman and Chapman, 1967, 1969) in which the authors describe a phenomenon they called "illusory correlation." Subsequently other articles appeared (Dowling & Graham, 1976; Golding & Rorer, 1972; Hamilton & Gifford, 1976; Hartsough, 1975; Rosen, 1975; Starr & Katkin, 1969; Tversky & Kahneman, 1973) purporting to confirm the reality of the phenomenon of "illusory correlation." This concept is now widely cited in the literature on clinical judgment, social judgment, person perception, and in personality and social psychology to support the general conclusion that human cognitive activity is flawed by certain biases (see, for example, Fischhoff, 1976; p. 431).

"Illusory correlation" was defined as "the report by observers of a correlation between two classes of events which, in reality, (a) are not correlated, or (b) are correlated to a lesser extent than reported, or (c) are correlated in the opposite direction from that which is reported" (Chapman & Chapman, 1967, p. 151). The Chapmans offered several examples of the phenomena they intend to include under the term "illusory correlation," among which was included Thorndike's halo effect.

In the first study, Chapman (1967) employed the paired-associates paradigm often used in memory research to show that the recalled frequency of the co-occurrence of pairs of simultaneously presented words with high associative value (e.g., bacon-eggs) was higher than the recalled frequency of neutral pairs of words (e.g., blossom-notebook). In the second and third studies (Chapman, & Chapman, 1967, 1969) extended their inquiry by examining the question of whether similar "illusory correlations" would be found in the cognitive activity of clinical psychologists in their diagnostic practice. The results of these studies had considerable impact, for the Chapmans reported that they had indeed discovered the same phenomena; they indicated that they had demonstrated "a source of massive systematic error in [clinicians'] observations of correlations between symptom statements (italics ours) and features of projective test protocols" (1969, p. 271), and that these "massive" errors were also based on the associative bond between sign and symptom. Following their study of clinicians they concluded:

Associatively based illusory correlation is a powerful bias. . . . Yet its influence is so unapparent that many practicing diagnosticians have overlooked it, and have substituted illusory correlates for valid correlates in their diagnostic practice." (1969, p. 280)

These charges of "massive systematic error" in diagnostic practice, apparently securely based on the empirical tested generalization of a basic research finding could hardly be ignored, and they were not. Subsequent researchers (indicated above), tested the generality of the Chapmans' results over other clinical tests by using the

same general research paradigm involving the recall of the frequency of clinical signs paired with symptom statements (broadly defined; e.g., homosexuality is regarded as a "symptom"). By 1976, Dowling and Graham could conclude that the presence of "illusory correlation" had been demonstrated with an objective clinical test (the MMPI) as well as with three projective tests (Draw-A-Person, the Rorschach, and the Incomplete Sentences Blank). The generalizability of the phenomenon of "illusory correlation" from word-pairs to inferences made in actual professional practice seemed firm.

"Illusory correlation," in short, received considerable attention, not only because it provided a rare and thus welcome example of generalization from basic research to clinical practice, but also because the generalization carried sharp practical implications for a professional activity seeking to confirm its social value; it provided further (and dramatic) empirical support for the hypothesis that clinical judgment has little if any validity, and showed, moreover, that the source of the lack of validity could be found in the well-known principle of associative bond, a mainstay of experimental psychology.

These conclusions regarding "massive" errors in clinical judgment, based on the discovery of "illusory correlation" are now fully established; Ross (1977), for example, refers to these studies as "classic demonstrations" (p. 198). They have been accepted by experimental psychologists; Tversky and Kahneman, 1973, for example, called the Chapman's results "ominous." And a leading clinical psychologist referred to them with the warning that "I do not believe that clinical

psychology will survive unless the effort to accumulate such knowledge becomes an intrinsic part of clinical psychology's role" (Rotter, 1973, pp. 320-321).

But these conclusions are unfounded. "Massive" errors in clinical judgment have not been shown to occur in the judgments of a single clinical psychologist (or other research subject). As will be shown below, the original study with word-pairs does not provide evidence for the presence of "illusory correlation" in clinical diagnosis. It was the entrapment by conventional experimental design and the failure to implement the concepts of representative design, ecological validity, and intra-ecological correlation that resulted in these erroneous, but highly important, scientific and professional conclusions.

Lack of Representativeness of Task Leads to Incorrect Generalizations about Psychological Processes

The paired associates paradigm used by the Chapmans (and subsequent researchers) is not representative of judgment tasks in general and clinical judgment tasks in particular. For clinical (and other) judgment tasks involve inferences from directly observed events to other events which are neither directly nor immediately observed. But the paired-associate paradigm does not require that the subject make an inference about indirectly observed events. The paired associate paradigm requires only the ability to learn and to recall the frequency of the co-occurrence of events, both of which have been directly observed. And that is what the experimenters told their subjects to do. Thus (p. 153), "The Ss were told before viewing the word-pairs that their task was to observe and report how often each word was paired with each other word."

To be sure, insofar as clinical (or other) judgment is inductive, it is based on experience, and of course, experience must somehow be encoded in memory. The learning and recall of co-occurrences may thus contribute to judgments of indirectly observed entities. But the paired associate paradigm does not require the subject to make a judgment about such inferred entities; it requires only that the subject learn and recall frequencies of paired events, both of which have been directly observed. The Chapmans, therefore, were studying paired associates learning and memory, not judgment, or inference.

Corroboration of this conclusion can be found in Tversky and Kahneman's (1973) replication of the Chapman's work in which they study "illusory correlation" as if it were a problem of memory, and reinterpret it in terms of the mnemonic concept of "availability." Wyer (1975, pp. 230-231) also makes the importance of the distinction clear by distinguishing "between cognitive organization, the manner in which previously-formed beliefs and attitudes are inter-related and stored in memory, and social inference, the process whereby several pieces of new information about an object are combined to form a new cognition about (or judgment of) this object" (italics in original). And Block (1977, p. 875) criticizes a similar study of "correlational bias" by Berman and Kenny (1976) for generalizing results obtained in a paired-associate task to observer's ratings of behavior.

In short, because the paired associate task used by the Chapmans did not evoke inference or judgment processes, their generalization of their results to clinical judgment processes was unwarranted.

They generalized their results to the class of relations they did not study; namely, the difference between the ecological validity of cues (their potential value in the inference process) and the utilization of cues (the extent to which the cues were actually used). In addition, they failed to draw the necessary distinction between objective intra-ecological correlations and subjective intra-ecological correlations among cues.

Distinctions between objective intra-ecological correlations between cues and subjective intra-ecological correlations between cues. Figure 1 illustrates the distinction between objective co-variations among cues in the environment and subjective co-variations among cues in a subject's cognitive system. Differences between objective co-variation and subjective co-variation among cues have been subject to study by those working in the Brunswikian framework (e.g., Armelius & Armelius, 1974, 1975, 1976; Brehmer, 1974, 1975; Knowles, Hammond, Stewart, & Summers, 1971, 1972;

-----  
Insert Figure 1 about here  
-----

Lindell, 1976; Lindell & Stewart, 1974; Mumpower & Hammond, 1974; Naylor & Schenck, 1968, among others). Figure 1 illustrates how such studies are carried out. Note that (a) each cue has a specific, quantified ecological validity (relation with the criterion variable) and (b) there are specific, quantified intra-ecological correlations between the cues. Thus, both classes of relations can be included and can be usefully differentiated in the proper context of inference and learning and memory, thus making it possible to determine (among other things) whether (a) subjects learn and retain knowledge about the degree and direction of the intra-ecological correlations between cues while (b) making judgments about the value or state of a criterion variable from cues with a specific ecological validity, (c) under various degrees of ecological validity and intra-ecological correlation between cues.

As an example of what might be learned in such Brunswikian studies, Mumpower and Hammond (1974) trained pairs of subjects to have highly different judgment policies. That is, in a two-cue inference task one person was trained to rely on (develop a high utilization of) cue No. 1 and to ignore cue No. 2. The other person was trained to rely on cue No. 2 and ignore cue No. 1. When required to make joint judgments in a task in which the cues were highly correlated, and without the benefit of information from the experimenter about their dissimilar judgment policies, the subjects did not discover that their judgment policies were in fact highly dissimilar--a case of "false agreement." But when the subjects were placed in a task in which the cues were uncorrelated with one another, they discovered

their fundamental disagreement rapidly. Thus it was learned that variations in intra-ecological correlations among (pairs of) directly observable cues can affect interpersonal conflict in the context of inferences about not directly observable events. (See also Brehmer, 1976, and Brehmer & Hammond, 1977, for a review of research on the effects of a variety of such formal task characteristics on interpersonal conflict.)

The Chapmans did not draw any of these distinctions, however; therefore, their conclusion regarding "illusory correlation" (a) unjustifiably generalizes results from one psychological process (learning and memory for the co-occurrence of events) to another (inference), (b) blurs conceptual distinctions already well-established between ecological validity, cue utilization and the intra-ecological correlation between cues, thus taking our capacity for conceptual and operational differentiation back to the period prior to 1943 (Brunswik, 1943) at least, and (c) over-generalizes conclusions drawn from studies of one class of relations (intra-ecological correlations) to a class of relations not studied (ecological cue validities and cue utilization).

In addition to these conceptual difficulties that lead to over- or mis-generalization, the work on "illusory correlation" also illustrates how failing to consider the methodological requirements of representative design leads to generalization in the wrong direction.

Representative Design, Statistical Tests, and Directional Errors of Generalization

Earlier we pointed out some examples of the failure to employ representative design in experiments designed to generalize results



over object-persons (see e.g., Rumenik et al., 1977). In the above examples the experimenters failed to sample the population of objects over which the generalization was intended (e.g., used an object-person sample of 1 or 2) while widely sampling the subject population. The subject sample  $n$  was then inserted in the statistical test (thus not only inflating the sample size, e.g., from  $n = 1$  or  $2$  to  $n = 300$ ) but achieving a generalization in the opposite direction from the generalization claimed. In the experiments testing the effect of the sex of the experimenter, for example, the generalization claimed is that the sex of the experimenters does (or does not) make a difference, but the generalization is based on the statistical test applied to a sample of subjects ( $n = 300$ ) not a sample of experimenters ( $n = 2$ ). The statistical test is made in the direction of subjects, but the generalization is made in the direction of objects.

It cannot be emphasized too strongly that these are not mere statistical errors to be corrected simply by consulting the proper statistician or textbook. Contrary to what is implied by Rumenik et al. (1977) and others, even expert statisticians who have cautioned against this error are susceptible to it. Rather, the error flows from the use of traditional methods based on the nomothetic, law-seeking, research paradigm in psychology. There is, therefore, ample reason to believe that this error will be slow to disappear. For although the number of studies described above (in which only one or two target persons are presented to hundreds of subjects) is diminishing because the use of persons as objects as well as subjects can make the directional error of

generalization seem simple and obvious, the directional error is not so obvious when inanimate objects are employed as the independent variable of an experiment. (For an important discussion of this problem in psycholinguistics, see Clark, 1973, who has argued for the necessity of (object) sampling of inanimate objects (words) in connection with research in studies of language; but see also Wike and Church, 1976.) The work on "illusory correlation" provides a useful example of why the concepts of object-sampling and representative design can be and must be applied to experiments involving inanimate objects as well as to those involving animate objects in order to avoid directional errors of generalization.

Subject and object sampling. In the first study in which the phenomenon of "illusory correlation" is reported, Chapman (1967) presents 6 word-pairs with high associative value (e.g., bacon-eggs) among other neutral word-pairs (e.g., blossoms-notebook) in such a manner that the word-pairs with high associative value appeared one-third of the time. The Ss were then asked to recall the frequency with which all the word-pairs occurred in the paired-associates learning task. As Chapman indicates, "the correct co-occurrence [of the high associative word pairs] in each case was. . .  $33\frac{1}{3}\%$ . It was predicted that for the six high associative pairs. . . the reported co-occurrence would be higher than this value. This excess over the correct value is the measure of illusory correlation" (p. 153).<sup>4</sup>

The first question is: why six word-pairs? Why not 10, or five, or four, or one word-pair? The answer is,

of course, that if too small a number of word-pairs were used, the results would not be convincing. Why not? Because words can vary in several dimensions (meaning, content, etc.) just as persons can vary in several dimensions, and, therefore, characteristics of word-pairs other than their associative value might account for the results. Consequently, the experimenter apparently somehow decided that six word-pairs would be an adequate number and sufficiently representative to provide generalizability. Thus it may be seen that the object side of this experiment was treated as the subject side of experiments was treated in the early 1900's, prior to the application of random sampling theory to subjects. Early psychophysicists, when not using a "good" subject, would employ an "adequate" number (e.g., 3, 4, 5) of subjects to insure the generalizability of their results. And although editors of psychological journals would today refuse to publish results based on such casual reasoning with regard to subject generalization, they do publish results based on such reasoning with regard to object generalization (including person-objects, e.g., Nisbett & Wilson, 1977, and inanimate objects, e.g., see any issue of any journal).

Indeed, it can be argued that the situation is worse today than it was prior to the use of sampling theory. For at that time the researchers merely over-generalized, that is, unjustifiably extended their results from an inadequate sample of subjects to populations of subjects. Today, however, the (mis)application of the logic of sampling theory misleads researchers into believing that they have tested the generalization intended, whereas, in fact, they have not tested it at all. In Chapman's study of "illusory correlation," for example, no statistical test

was made to determine whether the six word-pairs constituted an adequate sample size (irrespective of other sampling considerations) to test the hypothesis, nor is there any discussion of this problem.

But we know that some hypothesis was tested statistically, otherwise the work would not have been published, nor would it be cited as frequently as it has been. The hypothesis tested was "that for the six high associative pairs the reported co-occurrence would be higher than  $[33\frac{1}{3}\%]$ ." In order to do so, Chapman compared the mean responses (41.3%, 46.7%, 40.2%, 43.3%, 39.1%) of the subject sample for the reported co-occurrence of the six critical word pairs with the value  $33\frac{1}{3}\%$  "by means of a two-tailed, large sample [of subjects] t-test" and found the difference to be "significant for each of the six word pairs with high associative connections,  $z = 3.45$  or larger,  $p < .001$  in each case" (p. 153, italics ours). Thus, each of the six tests of significance was made over numerous subjects ( $n = 55, 49, \text{ and } 59$ ) but only one word-pair in each test. In accordance with conventional methodology, the conclusion was reached that since the difference was significant in each case, the phenomenon of "illusory correlation" had been observed (with the probability indicated).

It is precisely at this point that we address Bandura's complaint that arguments about the "ecological validity" of experiments are vague. We trust that readers will now see that it is wholly inappropriate to raise the question of the "ecological validity" of Chapman's 1967 study but that it is appropriate to raise the question of the representative design of that experiment. For while their study

applies sampling theory to the subjects of the experiment, it does not apply sampling theory to the objects (word-pairs) over which the generalization was intended.

Since six word-pairs of high associative value were used, however, we can address the question of whether any subject responded with a systematic bias in favor of the word-pairs with associative value? More specifically; is the response bias shown by any subject systematically different (at some level of confidence) from that expected by chance alone? Sampling theory applied to objects enables us to evaluate the question of whether six word-pairs are a sufficiently large sample to test that hypothesis. The weakest test is to ask only that a subject err on the side of overestimation by the minimal amount (34%). Even if that is all that is asked, however, a subject would have to overestimate on all six occasions to exceed the customary .05 probability level of significance (by a binomial test).

Without access to Chapman's 1967 data, one cannot ascertain whether any subject did overestimate on all six occasions. But it is doubtful that either Chapman or the journal editors would have accepted this minimal test because the small magnitude of the effect (1% overestimation) would have been considered trivial. How large an effect would not be considered trivial? With a sample size of six word-pairs, we can be confident (at the 5% level) that if any subject's mean estimate departed from the 33% value by 1.65 standard errors (1-tailed test) the subject was responding systematically to the highly associated six word-pairs. This means that only those

subjects whose mean estimate of the frequency of those word-pairs exceeded 65% could be considered to have become susceptible to "illusory correlation." Since the average estimate of Chapman's subjects in the first series was only 42.3% (by the third series it had dropped to 35.6%) it is extremely doubtful that the phenomenon of "illusory correlation" could have been demonstrated by any subjects. For if we set the criterion for evidence of systematic departure from  $33\frac{1}{3}\%$  at 65%, the lower (.05) confidence bound for 65% for a sample of, say, 50 persons, is 54%, thus far exceeding the mean of 42.3% obtained by Chapman on the first series. The conclusion that a mean estimate of 42.3% was a sufficiently large departure from  $33\frac{1}{3}\%$  to establish the generality of the phenomenon was, therefore, itself an illusion, because the statistical test was made in the direction opposite from the direction claimed; it was made in the direction of subjects, rather than objects (word-pairs).

But Chapman should not be singled out for criticism for using this mistaken methodology; it is common practice. Not only was this practice followed in every experiment conducted by Chapman, it was followed by every other researcher studying "illusory correlation," (but see Golding & Rorer, 1972, for a partial exception<sup>5</sup>).

In short, the phenomenon of "illusory correlation" may exist, but it has yet to be shown that any one person suffers from it. The unproven conclusions regarding "massive systematic errors," as well as the wide, uncritical acceptance of these conclusions, are due to entrapment by a research paradigm that commonly leads to errors of direction in generalization.

So pervasive is the entrapment by the conventional paradigm that even those who have urged that the conventional paradigm should be avoided fall victim to it. Tversky (1972), for example, advocates (as did Brunswik, 1952, 1956) that experiments should be conducted in such a manner that it can be ascertained whether any individual exhibits the phenomenon in question, thus:

most studies [of preference, choice, and judgment] report and analyze only group data. Unfortunately, group data usually do not permit adequate testing of theories of individual choice behavior because, in general, the compatibility of such data with theory is neither a necessary nor a sufficient condition for its validity. (p. 291)

Yet in Tversky's replication and elaboration of the original studies of "illusory correlation" (Tversky & Kahneman, 1973, pp. 225, 227) the conventional method of analysis of group, rather than individual, data was employed, resulting in the same error that Tversky cautioned against.<sup>6</sup> Unfortunately, such directional errors of generalization will almost certainly result in the rejection of the null hypothesis when it is true, thus filling the libraries with inconsequential results.<sup>7</sup>

#### Conclusion

In his eulogy of his colleague Brunswik, Edward Tolman (1956) wrote: "In the coming years, Egon Brunswik will hold an ever increasingly significant and important position in the history of psychology. . . . Those of us who knew and loved him can but be glad that such ever-greater recognition lies ahead, though we grieve.

that he did not live to see it happen." I would be less than honest if I denied that a similar personal commitment to Brunswik was present in my motivation to write this article.

Be that as it may, my purpose is also professional and impersonal. For the need for a psychology that can meet the rigorous demands of science and yet produce nontrivial results is clearly urgent, particularly in view of the growing conviction that the traditional approach is not serving us well, and that it is not serving society well. Not only does psychology need a scientific base that will provide the justification for professional practice, society needs psychological research that produces scientific generalizations that do not "decay," as Cronbach put it. Society also needs research results that will stand the test of generalization from laboratory to life.

These needs will be satisfied by preserving the character of the revolution in scientific psychology that was set in motion by Brunswik when he (a) urged the search for idiographic descriptions (models) of the behavior of individuals over situations, (b) showed the necessity for the use of representative design in order to test the validity and utility of such descriptions and (c) observed that these changes in direction would mean that the geographer's descriptions rather than the physicist's laws would constitute the proper aim for the work of the scientific psychologist.



## Reference Note

1. A report (dated August 12, 1977) of a readership survey of Journal of Personality and Social Psychology issued by the Chairman of the Publications and Communications Board of the American Psychological Association indicates that "almost half of the respondents (46%) viewed recent articles as less pertinent [to their interests than earlier articles]" and "close to half. . . wanted more field research" in contrast to the "4% [who] called for a greater emphasis on laboratory experiment."

References

- Armeliu, B., & Armeliu, K. The use of redundancy in multiple-cue judgments. Data from a suppressor variable task. American Journal of Psychology, 1974, 87, 385-391.
- Armeliu, B., & Armeliu, K. Note on detection of cue inter-correlation in multiple-cue probability learning. Scandinavian Journal of Psychology, 1975, 16, 37-41.
- Armeliu, K., & Armeliu, B. The effects of cue-criterion correlations, cue intercorrelations, and the sign of the cue intercorrelation on performance in suppressor variable tasks. Organizational Behavior and Human Performance, 1976, 17, 241-250.
- Aronson, E., Willerman, B., & Floyd, J. The effect of a pratfall on increasing personal attractiveness. Psychonomic Science, 1966, 4, 227-228.
- Aronson, E., & Carlsmith, J. Experimentation in social psychology. In G. Lindzey & E. Aronson (Eds.), The handbook of social psychology (2nd ed., Vol. 3). Reading, Mass.: Addison-Wesley Publishing Co., 1968.
- Bakan, D. A generalization of Sidman's results on group and individual functions, and a criterion. Psychological Bulletin, 1954, 51(1), 63-64.
- Bandura, A. On paradigms and recycled ideologies. Cognitive therapy and research, 1978, 2, 79-103.
- Berman, J., & Kenny, D. Correlational bias in observer ratings. Journal of Personality and Social Psychology, 1976, 34, 263-273.

- Block, J. Correlational bias in observer ratings: Another perspective on the Berman-Kenny study. Journal of Personality and Social Psychology, 1977, 35, 873-880.
- Brehmer, B. The effect of cue intercorrelation on interpersonal learning of probabilistic inference tasks. Organizational Behavior and Human Performance, 1974, 12, 397-412.
- Brehmer, B. Policy conflict and policy change as a function of task characteristics. IV. The effect of cue intercorrelations. Scandinavian Journal of Psychology, 1975, 16, 85-96.
- Brehmer, B. Social judgment theory and the analysis of interpersonal conflict. Psychological Bulletin, 1976, 83(6), 985-1003.
- Brehmer, B., & Hammond, K. R. Cognitive factors in interpersonal conflict. In D. Druckman (Ed.), Negotiations: Social-psychological perspectives. Beverly Hills, Calif.: Sage Publications, 1977.
- Bronfenbrenner, U. Toward an experimental ecology of human development. American Psychologist, 1977, 32(7), 513-531.
- Bruner, J. S., Goodnow, J. J., & Austin, G. A. A study of thinking. New York: Wiley, 1956.
- Brunswik, E. Organismic achievement and environmental probability. Psychological Review, 1943, 50, 255-272.
- Brunswik, E. The conceptual framework of psychology. In International Encyclopedia of Unified Science (Vol. 1, No. 10). Chicago: University of Chicago Press, 1952.
- Brunswik, E. Historical and thematic relations of psychology to other sciences. Science Monitor, 1956, 83, 151-161.

- Brunswik, E. Scope and aspects of the cognitive problem. In H. Gruber, R. Jessor, & K. R. Hammond (Eds.), Contemporary approaches to cognition: A symposium held at the University of Colorado. Cambridge: Harvard University Press, 1957.
- Chapman, L. J. Illusory correlation in observational report. Journal of Verbal Learning and Verbal Behavior, 1967, 6, 151-155.
- Chapman, L. J., & Chapman, J. P. Genesis of popular but erroneous psychodiagnostic observation. Journal of Abnormal Psychology, 1967, 72, 193-207.
- Chapman, L. J., & Chapman, J. P. Illusory correlation as an obstacle to the use of valid psychodiagnostic signs. Journal of Abnormal Psychology, 1969, 74, 271-280.
- Christensen, L. B. Experimental Methodology. Boston, Mass.: Allyn and Bacon, 1977.
- Clark, H. H. The language-as-fixed-effect fallacy. Journal of Verbal Learning and Verbal Behavior, 1973, 12, 335-359.
- Cronbach, L. Beyond the two disciplines of scientific psychology. American Psychologist, 1975, 30(2), 116-127.
- Crow, W. The need for representative design in studies of interpersonal perception. Journal of Consulting Psychology, 1957, 21, 321-325.
- Dowling, J., & Graham, J. Illusory correlation and the MMPI. Journal of Personality Assessment, 1976, 40, 531-538.
- Dudycha, A. L., & Naylor, J. C. The effect of variations in the Cue R matrix upon the obtained policy equation of judges. Educational and Psychological Measurement, 1966, 26, 583-603.

- Eaton, W. O., & Clore, G. L. Interracial imitation at a summer camp. Journal of Personality and Social Psychology, 1975, 32(6), 1099-1105.
- Einhorn, H. J. Expert measurement and mechanical combination. Organizational Behavior and Human Performance, 1972, 7, 86-106.
- Elms, A. C. The crisis of confidence in social psychology. American Psychologist, 1975, 30(10), 967-976.
- Fischhoff, B. Attribution theory and judgment under uncertainty. In J. Harvey, W. Ickes, & R. Kidd (Eds.), New directions in attribution research. Hillsdale, N.J.: Erlbaum Assoc., 1976.
- Fisher, R. A. The design of experiments (4th ed.). New York: Hafner, 1947.
- Frodi, A. On the elicitation and control of aggressive behavior. Göteborg Psychological Reports, 1974, 4(8), 16.
- Gergen, K. J. Social psychology as history. Journal of Personality and Social Psychology, 1973, 26, 309-320.
- Gergen, K. Social psychology, science and history. Personality and Social Psychology Bulletin, 1976, 2, 373-383.
- Gibson, J. Survival in a world of objects (A review of "Perception and the representative design of experiments"). Contemporary Psychology, 1957, 15, 694-703.
- Gillis, J. S. The effects of selected anti-psychotic drugs on objective task learning, and interpersonal learning with acute schizophrenics. In K. R. Hammond & C. R. B. Joyce (Eds.), Psychoactive drugs and social judgment: Theory and research. New York: Wiley, 1975.

- Gillis, J. S., & Schneider, C. The historical preconditions of representative design. In K. R. Hammond (Ed.), The psychology of Egon Brunswik. New York: Holt, Rinehart, and Winston, 1966.
- Goldberg, L. R. Man versus models of man: A rationale, plus some evidence for a method of improving on clinical inferences. Psychological Bulletin, 1970, 73(6), 422-432.
- Golding, S. L., & Rorer, L. G. Illusory correlation and subjective judgment. Journal of Abnormal Psychology, 1972, 80, 249-260.
- Graham, K. R. Psychological research: Controlled interpersonal interaction. New York: Brooks Cole, 1977.
- Greenwald, A. G. Within-subjects designs: To use or not to use? Psychological Bulletin, 1976, 83(2), 312-320.
- Hamilton, D. L., & Gifford, R. K. Illusory correlation in interpersonal perception: A cognitive basis of stereotypic judgments. Journal of Experimental Social Psychology, 1976, 12, 392-407.
- Hammond, K. R. Subject and object sampling--a note. Psychological Bulletin, 1948, 43, 38-48.
- Hammond, K. R. Representative vs. systematic design in clinical psychology. Psychological Bulletin, 1954, 51, 150-159.
- Hammond, K. R. Probabilistic functioning and the clinical method. Psychological Review, 1955, 62, 255-262.
- Hammond, K. R. The psychology of Egon Brunswik. New York: Holt, Rinehart, and Winston, 1966.
- Hammond, K. R. Egon Brunswik. In International encyclopedia of the social sciences, 1968, 156-158.

- Hammond, K. R., Rohrbaugh, J., Mumpower, J., & Adelman, L. Social judgment theory: Applications in policy formation. In M. F. Kaplan & S. Schwartz (Eds.), Human judgment and decision processes: Applications in problem settings. New York: Academic Press, 1977.
- Hammond, K. R., Stewart, T., Brehmer, B., & Steinmann, D. O. Social judgment theory. In M. F. Kaplan & S. Schwartz (Eds.), Human judgment and decision processes. New York: Academic Press, 1975.
- Hartsough, W. R. Illusory correlation and mediated association. Canadian Journal of Behavioral Science, 1975, 7, 151-154.
- Heider, F. The psychology of interpersonal relations. New York: Wiley, 1958.
- Hochberg, J. Perception: Toward the recovery of a definition. Psychological Review, 1956, 63(6), 400-405.
- Hochberg, J. Representative sampling and the purposes of perceptual research: Pictures of the world and the world of pictures. In K. R. Hammond (Ed.), The psychology of Egon Brunswik. New York: Holt, Rinehart, and Winston, 1966.
- Hull, C. The problem of intervening variables in molar behavior theory. Psychological Review, 1943, 50, 273-291.
- Jarvik, M. A functional view of memory. In K. R. Hammond (Ed.), The psychology of Egon Brunswik. New York: Holt, Rinehart, and Winston, 1966.
- Jenkins, J. J. Remember that old theory of memory? Well, forget it! American Psychologist, 1974, 29(11), 785-795.
- Keeley, S., & Doherty, M. Bayesian and regression modeling of graduate admission policy. Organizational Behavior and Human Performance, 1972, 8, 297-323.

- Knowles, B. A., Hammond, K. R., Stewart, T. R., & Summers, D. A.  
Positive and negative redundancy in multiple cue probability  
tasks. Journal of Experimental Psychology, 1971, 90, 157-159.
- Knowles, B. A., Hammond, K. R., Stewart, T. R., & Summers, D. A.  
Detection of redundancy in multiple-cue probability tasks.  
Journal of Experimental Psychology, 1972, 93, 425-427.
- Koch, S. Epilogue. In S. Koch (Ed.), Psychology: A study of  
science (Vol. 3). New York: McGraw-Hill, 1959.
- Kraut, R. Verbal and non-verbal cues in the perception of lying.  
Journal of Personality and Social Psychology, 1978, 36, 380-391.
- Kuhn, T. The structure of scientific revolutions. Chicago:  
University of Chicago Press, 1962.
- Leeper, R. W. A critical consideration of Egon Brunswik's probabilistic  
functionalism. In K. R. Hammond (Ed.), The psychology of Egon  
Brunswik. New York: Holt, Rinehart, and Winston, 1966.
- Lewin, K. Defining the field at a given time. Psychological Review,  
1943, 50, 292-310.
- Lindell, M. K. Cognitive and outcome feedback in multiple-cue  
probability learning tasks. Journal of Experimental Psychology,  
1976, 2(6), 739-745.
- Lindell, M. K., & Stewart, T. R. The effects of redundancy in  
multiple-cue probability learning. American Journal of  
Psychology, 1974, 87, 393-398.
- Loevinger, J. Psychological tests in the conceptual framework of  
psychology. In K. R. Hammond (Ed.), The psychology of Egon  
Brunswik. New York: Holt, Rinehart, and Winston, 1966.



- McGuire, W. J. The Yin and Yang of progress in social psychology: Seven koan. Journal of Personality and Social Psychology, 1973, 28, 446-456.
- Mumpower, J. L., & Hammond, K. R. Entangled task-dimensions: An impediment to interpersonal learning. Organizational Behavior and Human Performance, 1974, 11, 377-389.
- Murrell, G. Combination of evidence in perceptual judgment. In M. F. Kaplan & S. Schwartz (Eds.), Human judgment and decision processes: Applications in problem settings. New York: Academic Press, 1977.
- Naylor, J. C., & Schenck, E. A. The influence of cue redundancy upon the human inference process for tasks of varying degrees of predictability. Organizational Behavior and Human Performance, 1968, 3, 47-61.
- Neisser, U. Cognition and reality. San Francisco: W. H. Freeman & Co., 1976.
- Nisbett, R. E., & Wilson, T. D. The halo effect: Evidence for unconscious alternation of judgments. Journal of Personality and Social Psychology, 1977, 35(4), 250-256.
- Orne, M. T. Hypnosis, motivation, and ecological validity. In W. Arnold & M. Page (Eds.), Nebraska symposium on motivation. Lincoln, Neb.: University of Nebraska Press, 1970.
- Osgood, C. Discussion. In H. Gruber, R. Jessor, & K. Hammond (Eds.), Contemporary approaches to cognition: A symposium held at the University of Colorado. Cambridge: Harvard University Press, 1957.

- Parke, R. P. Social cues, social control and ecological validity. Merrill Palmer Quarterly, 1976, 22(2), 111-123.
- Postman, L., & Tolman, E. C. Brunswik's probabilistic functionalism. In S. Koch (Ed.), Psychology: A study of a science (Vol. 1). New York: McGraw-Hill, 1959.
- Rappoport, L., & Summers, D. A. Human judgment and social interaction. New York: Holt, Rinehart & Winston, 1973.
- Rosen, G. J. On the persistence of illusory correlation associated with the Rorschach. Journal of Abnormal Psychology, 1975, 84, 571-573.
- Ross, L. The intuitive psychologist and his shortcomings: Distortions in the attribution process. In L. Berkowitz (Ed.), Advances in experimental social psychology (Vol. 10). New York: Academic Press, 1977.
- Rotter, J. The future of clinical psychology. Journal of Clinical and Consulting Psychology, 1973, 40, 313-321.
- Rumenik, D., Capasso, D., & Hendrick, C. Experimenter sex effects in behavioral research. Psychological Bulletin, 1977, 84, 852-877.
- Selby, J. W., Calhoun, L. G., & Brock, T. A. Sex differences in the perception of rape victims. Personality and Social Psychology Bulletin, 1977, 3(3), 412-415.
- Sidman, M. A note on functional relations obtained from group data. Psychological Bulletin, 1952, 49, 263-269.
- Silverstein, C. H., & Strang, D. G. Seating position and interaction in triads: A field study. Sociometry, 1976, 39(2), 166-170.

- Slovic, P., & Lichtenstein, S. Comparison of Bayesian and regression approaches to the study of information processing in judgment. Organizational Behavior and Human Performance, 1971, 6, 649-744.
- Smedslund, J. Multiple probability learning. Oslo, Norway: Oslo University Press, 1955.
- Smith, M. B. Social psychology, science and history: So What? Personality and Social Psychology Bulletin, 1976, 2, 438-444.
- Starr, B. J., & Katkin, E. S. The clinician as an aberrant actuary: Illusory correlation in the Incomplete Sentences Blank. Journal of Abnormal Psychology, 1969, 74, 670-675.
- Steinmann, D., & Doherty, M. Lens model analysis of bookbag and poker chip experiment: A methodological note. Organizational Behavior and Human Performance, 1972, 8, 450-455.
- Stewart, T. R. Components of correlation and extensions of the lens model equation. Psychometrika, 1976, 41, 101-120.
- Tajfel, H. Social and cultural factors in perception. In G. Lindzey and E. Aronson (Eds.), The handbook of social psychology (2nd ed., Vol. 3). Reading, Mass.: Addison-Wesley Publishing Co., 1968.
- Thorndike, E. L. A constant error in psychological ratings. Journal of Applied Psychology, 1920, 4, 25-29.
- Tolman, E. C. Eulogy: Egon Brunswik: 1903-1955. American Journal of Psychology, 1956, 69, 315-342. Reprinted in K. R. Hammond (Ed.), The psychology of Egon Brunswik. New York: Holt, Rinehart, & Winston, 1966.

- Tversky, A. Elimination by aspects: A theory of choice. Psychological Review, 1972, 79, 281-299.
- Tversky, A., & Kahneman, D. Availability: A heuristic for judging frequency and probability. Cognitive Psychology, 1973, 5, 207-232.
- Wallace, H. A. What is in the corn judge's mind? Journal of American Society of Agronomy, 1923, 15, 300-304.
- Wike, E., & Church, J. Comments on Clark's "The language-as-fixed-effects fallacy." Journal of Verbal Learning and Verbal Behavior, 1976, 15, 249-255.
- Wyer, R. S., Jr. The role of probabilistic and syllogistic reasoning in cognitive organization and social inference. In M. F. Kaplan & S. Schwartz (Eds.), Human judgment and decision processes. New York: Academic Press, 1975.

## Footnotes

1. Preparation of this article was assisted by grants from the National Science Foundation (BNS-76-09560) and the Office of Naval Research (N00014-77-0336). Reprints may be obtained by writing to the author at the Center for Research on Judgment and Policy, Institute of Behavioral Science, University of Colorado, Boulder, Colorado, 80309.

2. This article was initiated as a result of conversations with Berndt Brehmer, to whom I am grateful for his many insights and suggestions that span many years. I also thank others, too many to name, for their criticism of early drafts.

3. It is interesting to observe that R. A. Fisher, the inventor of analysis of variance techniques, was keenly aware of the limitations of factorial designs beyond their obvious utility for the discovery of specific, nongeneralizable effects such as those sought in applied research. For example, "The exact standardization of experimental condition, which is often thoughtlessly advocated. . . , always carries with it the real disadvantage that [it]. . . supplies direct information only in respect of the narrow range of conditions achieved by standardization. Standardization, therefore, weakens rather than strengthens our ground for inferring a like result, when, as is invariably the case in practice, these conditions are somewhat varied" (1947, p. 97).

4. The unfortunate and unjustified extension of co-occurrence within the paired associate paradigm to bring it under the concept of "correlation" should not go un-noticed. Correlation, of course, refers to the co-variation between dimensions over a range of values, not simply paired association. Co-variation did not appear in the first study of "illusory correlation," nor is it typical of later studies.

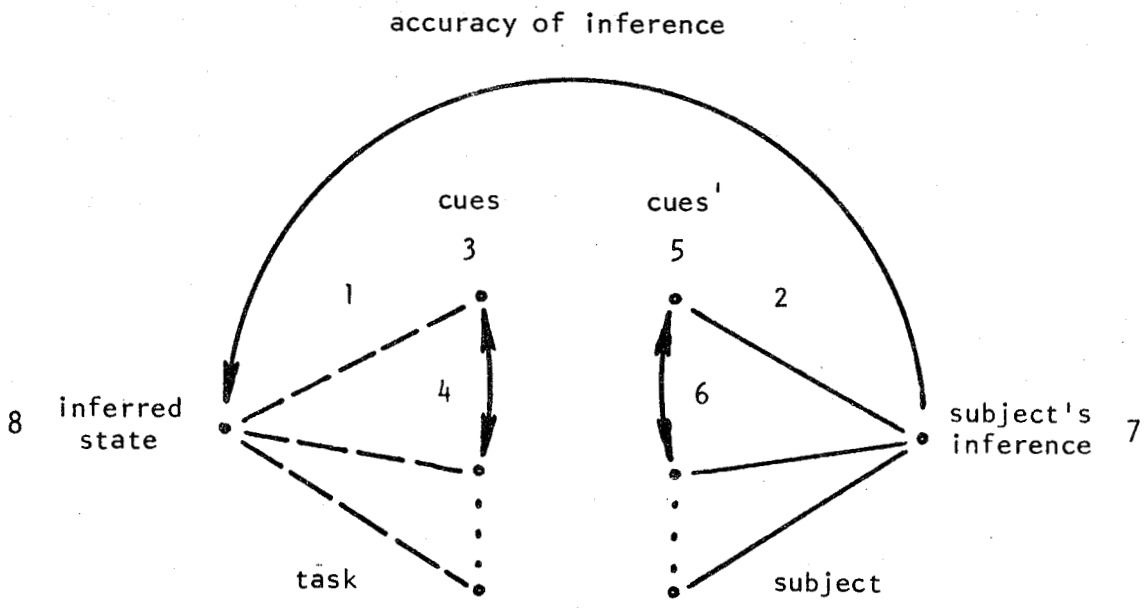
5. Golding and Rorer (1972) are a "partial exception" because although they used the number of subjects in their principal tests of differences between group means in the same manner as the other researchers, they also included some tests of individual subjects' behavior, apparently because of their knowledge of the Brunswikian tradition. These tests provided only weak evidence for errors in associations.

6. In their replication of the Chapman's study of "illusory correlation" with word-pairs Tversky and Kahneman (1973) use a larger sample of words (10 word-pairs of high associative value and 10 word-pairs of neutral associative value) than the Chapmans but, like Chapman and others, change the direction of the generalization by inserting in their test the number ( $n = 30$ ) of subjects (p. 225). In a somewhat different test (p. 227) eight word-pairs were used in one set, eight in the other, but the number of subjects ( $N = 62$ ,  $N = 68$ ,  $N = 73$ , in three different studies) was employed in the statistical test, again changing the direction of the test.

7. See Bakan (1954), also Sidman (1952), for supporting mathematical developments of this argument applied to the practice of generalizing from group averages to individual preference, particularly with regard to learning curves.

Figure Caption

Figure 1. Studies of illusory correlation compare the objective frequency of directly observable co-occurring events (4) with the subjects' recall of the frequencies of such events (6). Judgment, on the other hand, involves inferences (7) from the observed events (3) through their psychological transformation (5) to the unobserved state (8). If an independent measure of (8) is available, then the accuracy of the judgment may be determined by comparing (7) with (8).



- 1 ecological validity
- 2 cue utilization
- 3 objective values of cues
- 4 objective inter-correlations among cues
- 5 subjective values of cues
- 6 subjective inter-correlations among cues
- 7 judgment
- 8 inferred state

Figure 1