Evaluation and Program Planning, Vol. 12, pp. 355-366, 1989
Printed in the USA.

0149-7189/89 $3.00 + .00

OUTCOME PATTERN MATCHING AND PROGRAM THEORY

WILLIAM M.K. TROCHIM

Cornell University

ABSTRACT

Pattern matching is presented as a general framework which can guide the use of theory within program evaluation. Pattern matching minimally involves the specification of a theoretical pattern, the acquisition of an observed pattern, and an attempt to match these two. Pattern matching logic assumes that more complex theoretical patterns, if corroborated, provide a stronger basis for valid inference. Pattern matches in program evaluation can be divided into two types: process pattern matches *which assess the construct validity of the program, participants, or measures, and* outcome pattern matches *which assess the causal hypothesis and address the traditional concerns of internal and external validity. Each of the three types of process pattern matches can be further divided into characteristic pattern matches (which examine the interrelationships between key characteristics across programs, participants, or measures) or object pattern matches (which view interrelationships between programs, participants, or measures based on their overall degree of similarity). Outcome pattern matching can be accomplished for any process pattern match by examining outcomes across programs, participants, or measures viewed either in terms of their characteristics or as molar objects. Hypothetical examples of pattern matching in program outcome evaluation contexts are presented along with consideration of the value of pattern matching for theory-based research.*

Over the past few years there has been a concerted effort within applied research circles to reemphasize the importance of theory in the research process. Chen and Rossi (1983, 1987) have played a major role in this effort as have Bickman (1986) and others. Chen and Rossi (1987) have taken as their starting point the popularity of theories of validity expounded by Campbell and Stanley (1963, 1966), Cook and Campbell (1979), Cronbach (1982), and Cronbach et al. (1980) which, they argue, have emphasized some validity types over others and have not adequately encouraged the potential of a theory-driven approach.

In one sense, what they are advocating is not new. Almost all researchers, including the ones they criticize, readily recognize the importance of theory in research. But in another sense Chen and Rossi (1987) have provided a great service in criticizing a research paradigm which has encouraged an experimental perspective at the expense of richer theoretical elucidation. However, while they argue for the greater use of theory, they do little to explain *how*, once a theory has been developed, it can be used within the research process. If theory is to become a more integral part of the research act, it is essential that different ways to use theory in research be better described.

This paper discusses how theory can be used in outcome evaluation by emphasizing the central role of pattern matching in research thinking. The idea of pattern matching stems from a long tradition of research thinking which, ironically enough, originated with the same persons most heavily criticized by Chen and Rossi (1987). It is seen perhaps most early in the efforts by Cronbach and Meehl (1955) to address the construct validity of measures by arguing for the development of a "nomological network" which explicitly links theoretical suppositions and constructs with operationalizations

This paper is an abridged version of a longer treatment of pattern matching which presents greater detail on the process pattern matches. Copies of the longer paper may be obtained directly from the author.

I wish to express my personal thanks to Thomas Kellaghan, Mary Rohan, and all of the staff of the Educational Research Centre at St. Patrick's College, Dublin, Ireland for graciously providing the environment within which it was possible to write this paper. I also wish to thank Mel Mark, Val Caracelli, Jules Marquart, and the students in my graduate research methods class for their constructive comments on drafts of this paper, even if I didn't always heed their advice.

Requests for reprints should be sent to William M.K. Trochim, Assistant Professor, Department of Human Services, N132 MVR Hall, Cornell University, Ithaca, NY 14853.

as part of the inferential process. Pattern matching, as described here, was an integral part of the thinking behind the "multitrait-multimethod" approach to construct validity set forth by Campbell and Fiske (1959) and later was described in greater detail in the paper "Pattern matching as an essential in distal knowing" (Campbell, 1966). Pattern matching principles are also implied in the more recent statement on causal inference by Einhorn and Hogarth (1986) who argue for the importance of looking at putative causes within their broader causal field or context.

As an explicit concept useful for research, however, pattern matching has only recently begun to be investigated. Trochim (1985) explained the role of pattern matching in research and provided examples of how such an approach strengthens traditional approaches to validity. Other researchers (Cordray, 1986; Mark, 1986) have begun incorporating pattern matching ideas more directly into their writings about the current state of quasi-experimental thinking. Recent work on the use of pattern matching in research is described in Caracelli (1989), Davis (1989), Dumont (1989), Galvin (1989), Keith (1989), and Marquart (1988, 1989).

The intent of this paper is to show that pattern matching is a useful mechanism for utilizing theory in outcome evaluations. A firm knowledge of pattern matching principles will guide the researcher in developing better theory and integrating that theory into the research process. This paper first presents the general notion of pattern matching and describes the major types of process and outcome pattern matches. These ideas are then illustrated with several outcome pattern matching examples (space considerations preclude presentations of examples of program, participant, or measurement pattern matching). Finally some general conclusions are drawn about the implications of a pattern matching approach for program theory and program evaluation.

THE THEORY OF PATTERN MATCHING

A pattern is any arrangement of objects or entities. The term "arrangement" is used here to indicate that a pattern is by definition nonrandom and at least potentially describable. All theories imply some pattern, but theories and patterns are not the same thing. In general, a theory postulates structural relationships between key constructs. The theory can be used as the basis for *generating* patterns of predictions. For instance, $E = MC^2$ can be considered a theoretical formulation. A pattern of expectations can be developed from this formula by generating predicted values for one of these variables given fixed values of the others. Not all theories are stated in mathematical form, especially in applied social research, but all theories provide information which enables the generation of patterns of predictions.

Pattern matching always involves an attempt to link two patterns where one is a theoretical pattern and the other is an observed or operational one. The basic idea of pattern matching is illustrated in Figure 1. The top part of the figure shows the realm of theory. The theory might originate from a formal tradition of theorizing, might be the ideas or "hunches" of the investigator, or might arise from some combination of these. The conceptualization task involves the translation of these ideas into a specifiable theoretical pattern indicated by the top oval in the figure. The bottom part of the figure indicates the realm of observation. This is broadly meant to include direct observation in the form of impressions, field notes, and the like, as well as more formal objective measures. The collection or organization of relevant operationalizations (i.e., relevant to the theoretical pattern) is termed the observational pattern and is indicated by the lower oval in the figure. The inferential task involves the attempt to relate, link, or match these two patterns as indicated by the double arrow in the center of the figure. To the extent that the patterns match, one can conclude that the theory and any other

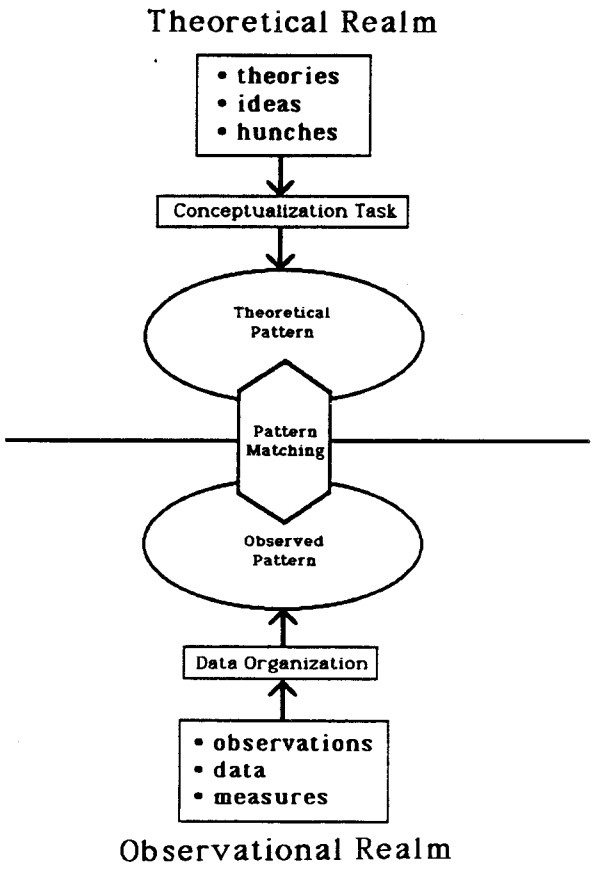

Figure 1. The basic pattern matching model.

theories which might predict the same observed pattern receive support.

It is important to demonstrate that there are no plausible alternative theories that account for the observed pattern and this task is made much easier when the theoretical pattern of interest is a unique one. In effect, a more complex theoretical pattern is like a unique fingerprint which one is seeking in the observed pattern. With more complex theoretical patterns it is usually more difficult to construe sensible alternative patterns which would also predict the same result. To make this more concrete, consider a theoretical outcome pattern for an educational program evaluation where it is hypothesized that the program will have its greatest effect on measures of immediate recall of course-related information, less of an effect on related attitudes, and the smallest effect on behavioral changes. If this pattern of outcomes is obtained, one might be tempted to conclude that the program caused the observed pattern to occur. However, there may be a plausible alternative explanation for this observed pattern. For instance, it may be that recall measures are more reliable than attitudinal ones which, in turn, are more reliable than behavioral ones. The observed pattern in this instance may be due to the pattern of reliability across measures rather than to the program. In this case, one would have to rule out the reliability-based explanation (perhaps by examining reliabilities to see if they are in fact distributed in this manner or by incorporating more measures of each type with differing reliabilities) before concluding that the program caused the outcome pattern. To the extent that theoretical and observed patterns do not match, the theory may be incorrect or poorly formulated, the observations may be inappropriate or inaccurate, or some combination of both states may exist.

All research employs pattern matching principles, although this is seldom done consciously. In the traditional two-group experimental context, for instance, the typical theoretical outcome pattern is the hypothesis that there will be a significant difference between treated and untreated groups. The observed outcome pattern might consist of the averages for the two groups on one or more measures. The pattern match is accomplished by a test of significance such as the *t*-test or ANOVA. In survey research, pattern matching forms the basis of generalizations across different concepts or population subgroups. In qualitative research pattern matching lies at the heart of any attempt to conduct thematic analyses.

While current research methods can be described in pattern matching terms, the idea of pattern matching implies more, and suggests how one might improve on these current methods. Specifically, pattern matching implies that *more complex patterns, if matched, yield greater validity for the theory*. For instance, if we complicate the two-group experimental case by introducing different treatment groups with different dosage levels *and* if the observed pattern matches the dosage-based expectations, we obtain a stronger corroboration of the theory. When viewed in this way we see the experimental paradigm criticized by Chen and Rossi (1987) as weak not for its *lack* of theory but rather for its reliance on simplistic theory. A pattern match in the traditional experiment provides evidence in favor of the theory only as long as there are no other plausible theoretical patterns which yield the same expected outcome pattern. In randomized experiments, the act of random assignment helps to equate groups probabilistically and minimize the likelihood of many competing theoretical patterns. The problem is much more acute, however, in quasi-experimental research where other factors such as selection may readily account for the same outcome pattern of differences between groups. A more "complex" dosage-based study, whether experimental or quasi-experimental, will in general reduce the number of competing patterns because it will be more difficult to construct plausible competing patterns which match the more "complex" pattern of results. Furthermore, the above discussion is framed in terms of looking only for an outcome pattern match. Even randomized experiments, however, can benefit from utilizing other relevant pattern matches (e.g., program, measurement, participant).

Pattern matching does not differ fundamentally from traditional hypothesis testing and model building approaches. A theoretical pattern is a hypothesis about what is expected in the data. The observed pattern consists of the data which are used to examine the theoretical model. The major differences between pattern matching and more traditional hypothesis testing approaches are that pattern matching encourages the use of more complex or detailed hypotheses and treats the observations from a multivariate rather than a univariate perspective.

The emphasis in this paper is on the use of pattern matching for outcome program evaluations or causal assessments. It should be noted, however, that pattern matching is not limited only to these cases. Pattern matching is an essential part of any process or formative evaluation in that it enables us to examine the construct validity of the program, sampling, or measures. In fact, pattern matching principles are even sensibly extended beyond the realm of applied social research or program evaluation. Trochim and Linton (1986) have suggested that the development of patterns is fundamental to both strategic and operational planning efforts. This can be seen most clearly in the discussion of program pattern matching which is intimately related to the program planning task.

There are several issues related to the theory of pattern matching which deserve mention here but which re-

quire more extensive treatment than this paper can provide. First, there is the question of how one can best develop the theoretical pattern for a given study. This is the conceptualization issue and has been discussed in Trochim and Linton (1986), Trochim (1985), and Trochim (1989). In brief, there is no one correct way to accomplish the conceptualization task and there is no one correct form which a theoretical pattern must take. Patterns may be verbal in nature, be a collection of mathematical formulae, or consist of a pictorial representation. There are many approaches which can be used to help develop theoretical patterns and conceptualization would be greatly enhanced by continued research and exploration of these approaches. This paper tends to offer pattern matching examples which are based on the conceptualization methods outlined in Trochim and Linton (1986), and Trochim (1985, 1989) which involve input from multiple relevant constituency groups in a research project and the use of multidimensional scaling and cluster analysis to generate two-dimensional pictorial concept maps which constitute theoretical patterns. This emphasis should be taken less as an advocated approach than as a reflection of the major experience of the author.

Second, when using pattern matching it is important to be aware of the "level of generalization" (Mark, 1986) at which one is operating. A theory can be stated at different levels of generality. For instance, in devising an educational program, one could develop a theoretical pattern which is specific to the study (e.g., the School X preschool program), a pattern which generalizes the specific program to other contexts (e.g., the School X approach to preschool education), or a pattern which treats the program construct very broadly (e.g., preschool education in general). The level of generalization which one chooses for any theoretical pattern should depend on the purposes of the study and the generalizability domain which is desired. How one chooses the level of generalization is an important conceptualization issue but is not formally considered here.

Third, this presentation of pattern matching tends to assume that theoretical patterns are developed prior to observation. A pattern match will tend to be most convincing when the theory is developed without knowledge of the observed pattern. However, pattern matching can and should be employed after the fact as part of a fuller exploration of the data or as the basis for the development of theoretical patterns which might be explored in subsequent studies. Pattern matching might be especially valuable in secondary reanalyses of databases that were previously analyzed with a more traditional approach.

Fourth, a pattern matching approach implies a different view of data than is common in most research. Specifically, it treats relevant data about programs, measures, participants, or outcomes as patterns or as a whole rather than just as a collection of individual measures or observations. Thus in a study which collects 100 different outcome measures, it is the *pattern* of outcome across these measures which is emphasized, not the outcome reflected by any one measure or aggregate index of the measures.

Finally, a key issue in pattern matching concerns the procedures which will be used to provide evidence for a match. In some cases, statistical analyses readily exist which are appropriate, but in most instances, acceptable statistical techniques are lacking or will have to be devised by the researcher. Further research into appropriate analytic procedures is needed although even straightforward visual approaches to pattern matching may be sufficient in many research settings.

TYPES OF PATTERN MATCHES IN RESEARCH

The major types of pattern matches (i.e., PMs)[1] for a typical program evaluation are described in greater detail in Figure 2. The figure makes a distinction between three *process pattern matches*—for the program, measures, and participants—and *outcome pattern matches*. Each process PM can be divided into two types—a characteristic-oriented and an object-oriented PM. A *characteristic PM* involves consideration of the characteristics of that part of the process. For instance, a program-characteristic PM would involve looking at the characteristics which define the program within some context (e.g., program content, setting, duration, method of delivery, materials). The *object PM* involves consideration of that part of the process as intact molar units. Thus, the program-object PM would involve examination of the interrelationships between a set of intact programs, where each program is treated as a separate molar object in and of itself. For each process PM which is conducted it is possible to implement a separate outcome PM. An outcome PM simply involves looking at how outcomes are distributed across programs, measures, or participants, considered either in terms of their characteristics or as objects.

In any program evaluation one or more of these PMs may be investigated. However, each PM imposes requirements upon the evaluation which may or may not be achieved in a specific study. For instance, one cannot conduct a program-object PM unless there are multiple programs being considered simultaneously. The reader should note that the assumption here is that process and outcome PMs already exist in every outcome evaluation or causal hypothesis test although they may not be described as such. Furthermore it is argued that

[1]The initials 'PM' will be used in the remainder of this paper to refer to specific pattern matches.

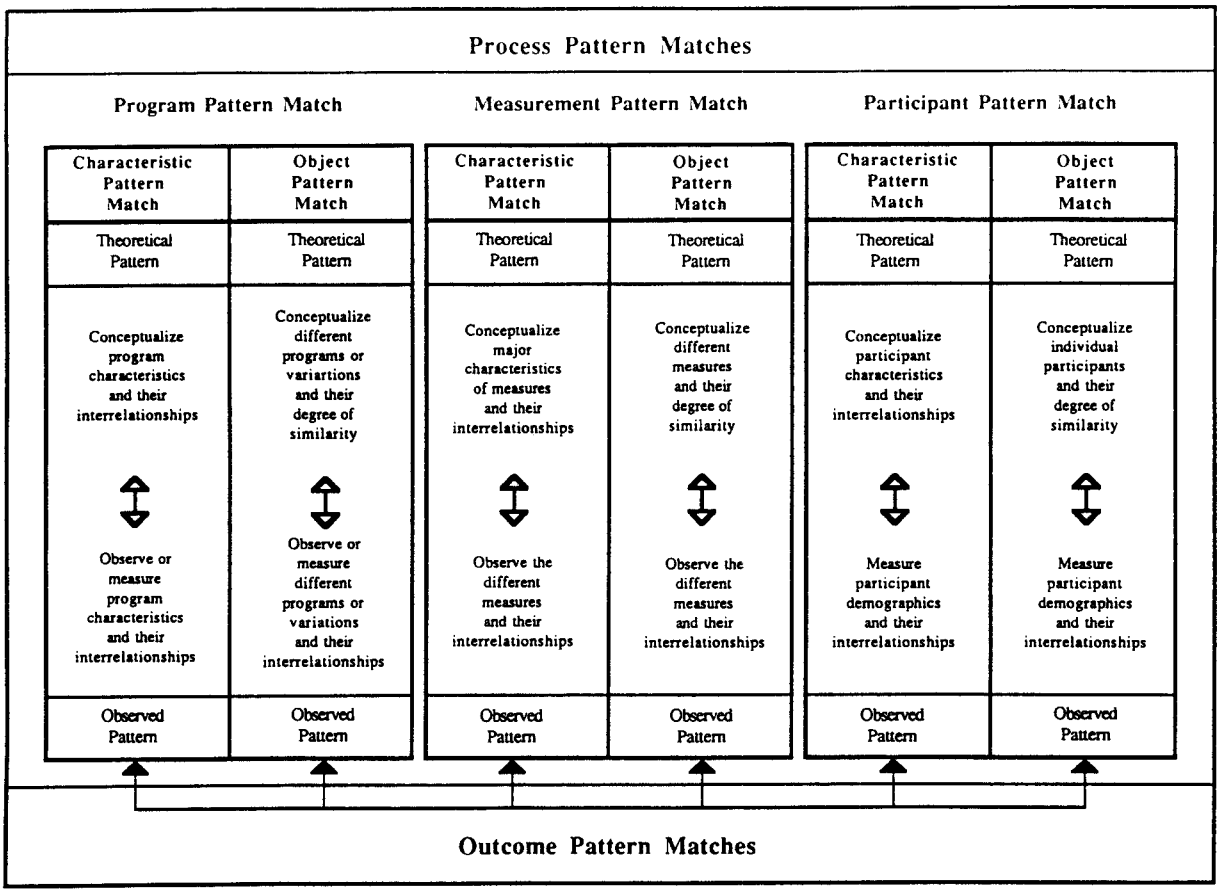

Figure 2. Types of pattern matches in an outcome evaluation.

existing research can be improved by identifying implicit PMs and specifying them more clearly.

For each PM, one identifies a theoretical and an observed pattern. Theoretical patterns imply more than just an arbitrary listing of the relevant concepts. Instead, a theoretical pattern involves describing the major relevant concepts *and their theoretical interrelationships*. To develop the theoretical-characteristic pattern for programs, one needs to describe the major characteristics of the programs and state how these characteristics are thought to be interrelated. To develop the theoretical-object pattern for programs, one needs to define different programs and state the expected degree of similarity between them.

Analogously, an observed pattern describes more than simply a collection of observations. Instead, an observed pattern should include observations *and their interrelationships*. To develop the observed-characteristic pattern for programs, one might use observations of the major characteristics of the programs and how these observations are interrelated. Similarly, to develop the observed-object pattern for programs, one might use observations or ratings of the degree to which programs are similar in practice.

The different types of PMs are relatable to commonly held taxonomies of validity types. In general, the three process-related PMs address issues of the construct validity of the program, measures, and sampling, respectively. Outcome PMs address both internal validity and the generalizability (external validity) of results. Because outcome PMs are always based on one of the three process PMs, one can conclude that a better job of delineating construct validity (i.e., process) will in general contribute to better internal and external validity. Trochim (1985) discusses the intimate relationship between construct and internal validity which is implied by a pattern matching perspective.

OUTCOME PATTERN MATCHING

In this section we focus our attention specifically on using pattern matching to assess the outcome of a program—the causal hypothesis. The assessment of the outcome of a study can be done in reference to one or all of the process pattern matches (i.e., program, measurement, and participant). Thus, in any study one

could assess the outcome of the program in relation to program context, participants, and/or measures.

All outcome pattern matching requires a theoretical pattern of expected outcomes, an observed pattern of effects, and an attempt to match the two. If one has already constructed pattern matches for the measures, participants, or programs, then the theories implicit in those pattern matches can help in generating the theoretical outcomes, and in assessing the effect of the program and the generalizability of that effect.

We can perhaps most sensibly begin to understand outcome pattern matching by linking outcomes to the measurement-object pattern match. To develop the theoretical outcome pattern in this case we might begin by considering the measurement in a hypothetical study. For example, imagine a study which involves a set of ten measures (two each of five separate constructs of skills). We might first state our expectations about the *relative* effects of the program of interest on each of the ten measures by ordering the ten measures from those which we believe will be most to least affected by the program. The observed outcome pattern could consist of simple tests of the significance of pre-post gains (e.g., *t*-tests) on the ten measures. The match would involve looking for a correspondence between the theoretical expectations and the obtained results. The procedure could be improved by describing our theoretical expectations at an interval level by having a group of judges or research participants estimate on some scale their expectations about how much the program will affect each measure. This idea is illustrated in Figure 3. The left side of the figure consists of the theoretical outcome pattern across measures and is based on a hypothetical average of judges' ratings of expected outcomes (where 1=negligible effect and 5=maximum effect). Measure m_1 is expected to show the greatest effect while measure m_{10} is expected to show the least. The right side of the figure graphs the *t*-values for each of the ten measures. Lines are drawn between the corresponding expectation and *t*-value for each of the measures. A perfect pattern match would be indicated by a perfect correlation between expectations and outcomes. In the figure, while the correlation is not perfect, it is strong and the general ordering of expected outcomes receives some confirmation. There are only two crossovers and even in these cases they are between adjacent measures.

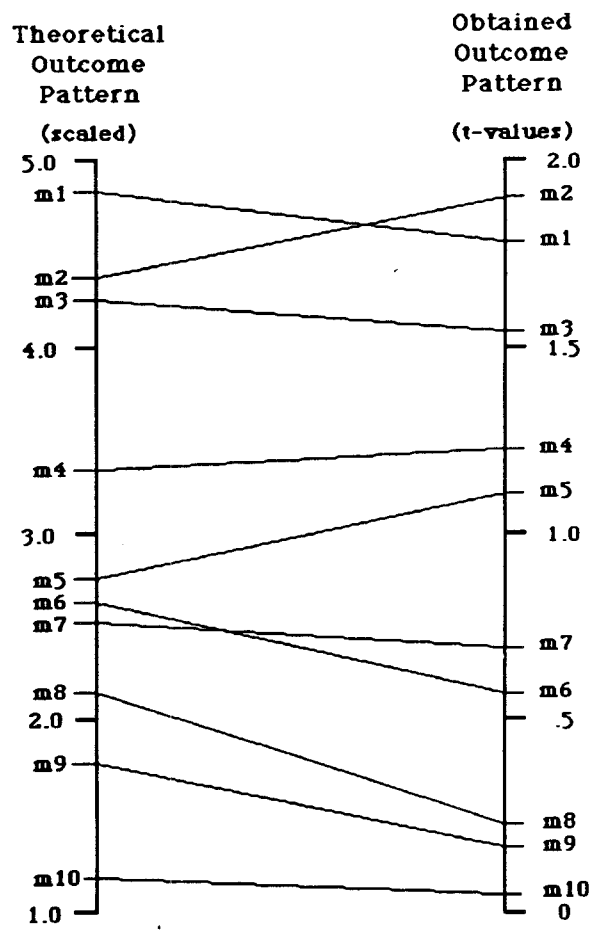

Figure 3. Hypothetical outcome pattern match using a measurement object pattern.

A number of features are important here. Virtually none of the individual *t*-values in this hypothetical example is significant at the traditional .05 level (i.e., $t>1.96$). Ordinarily, if one were interpreting these ten *t*-tests the typical conclusion would be that the program has had no significant effect. Nevertheless, the correspondence between the expectations and the obtained *t*-values is striking and it would be extremely difficult to pose an alternative explanation for this correspondence other than that the program has had an effect. In an example like this, the use of pattern matching might considerably alter the conclusions which one would reach. The match could be assessed using a simple correlation between theoretical expectations and observed *t*-values, although this would probably underestimate the strength of the match because it would be based on a sample size equal to the number of tests regardless of how many participants the individual *t*-tests were based on.

We could extend this example further by overlaying outcome estimates onto a measurement object pattern map. In this example, the measurement pattern map might be obtained by conducting a multidimensional scaling analysis of judges estimates of similarity between the different possible measures for a study. The multidimensional scaling coordinates can then be mapped and if we overlay the *t*-values onto this map pattern we might obtain something like the hypothetical graph shown in Figure 4. The theoretical pattern is shown in the top half of the figure and would consist of our expectations that arithmetic measures should be most affected and spelling ones least affected by the

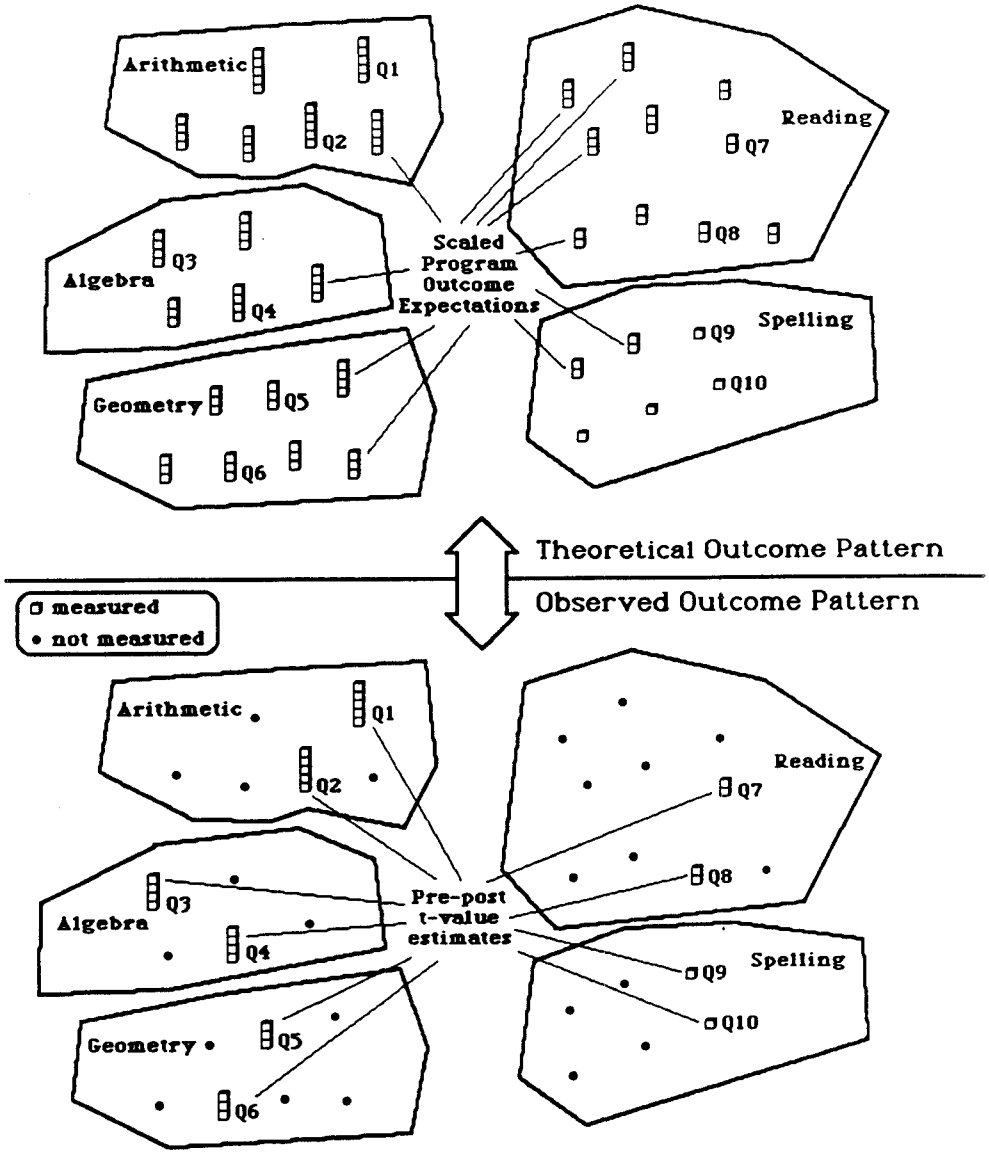

Figure 4. Hypothetical outcome pattern match using a measurement object pattern.

program. Higher bars on the top half of the figure indicate that judges expect greater program effects for that measure. The hypothetical observed results as indicated on *t*-tests are shown in the bottom half of the figure for only the concepts that were measured in the study (i.e., a larger number of skill items were conceptualized as potentially measureable, but only two items were selected per construct for the final study. The rest are shown as dots on the figure). The pattern of results clearly indicates that for the arithmetic program being studied the results were strongest for the arithmetic measures and weakest for the spelling ones as expected. Thus we can use the theory which is implicit in the measurement interrelationships to assess the overall effect of the program irrespective of the statistical significance of any of the individual *t*-tests. We are, in a sense, placing outcomes within a conceptual context. When we obtain a pattern match in this manner even with nonsignificant *t*-values, we would most likely conclude that while the statistical power of the significance tests may be lower than desired, the program had a detectable *pattern* of effect. Of course, any match which is obtained would still need to be evaluated in terms of practical significance.

By using the measurement structure shown in Figure 4, we can also reach some conclusions about the generalizability of the effect across measurement constructs. For instance, we might conclude that the program effect on arithmetic questions which were *not* included as measures (shown as dots in the bottom half of the figure) would be closer to the effect obtained on the two measured arithmetic questions than on spelling

or geometry ones. We can reach this conclusion only because we have already developed a theory or construct pattern which, through the measurement PM, has confirmed that the nonincluded arithmetic questions are more similar to the included ones than to the spelling or geometry ones.

A second type of outcome PM can be constructed on the basis of a program PM. For instance, if we consider a hypothetical study of three programs or program variations (conceptualized within a broader context of many program types), it would be possible for us to map program effects for any one measure onto the three programs shown in the figure as indicated in Figure 5. The top part of the figure shows the expected outcomes across programs. This map might be generated by performing multidimensional scaling of judges' estimates of the degree of similarity between all potential program variations. The coordinates could then be graphed and average estimates of expected program effect for each program type could be overlayed onto this graph as shown in the top part of the figure. Programs are located on the map on the basis of their shared characteristics. For instance, Programs 1 and 2 might emphasize primarily arithmetic while Program 3 focuses on some other skill (e.g., spelling). Our expectation would be that Program 1 should show the highest effect on measures most related to the nature of that program (e.g., the arithmetic program). Program 3 would be expected to show little or no effect on arithmetic. Furthermore we would expect that the results for Program 2

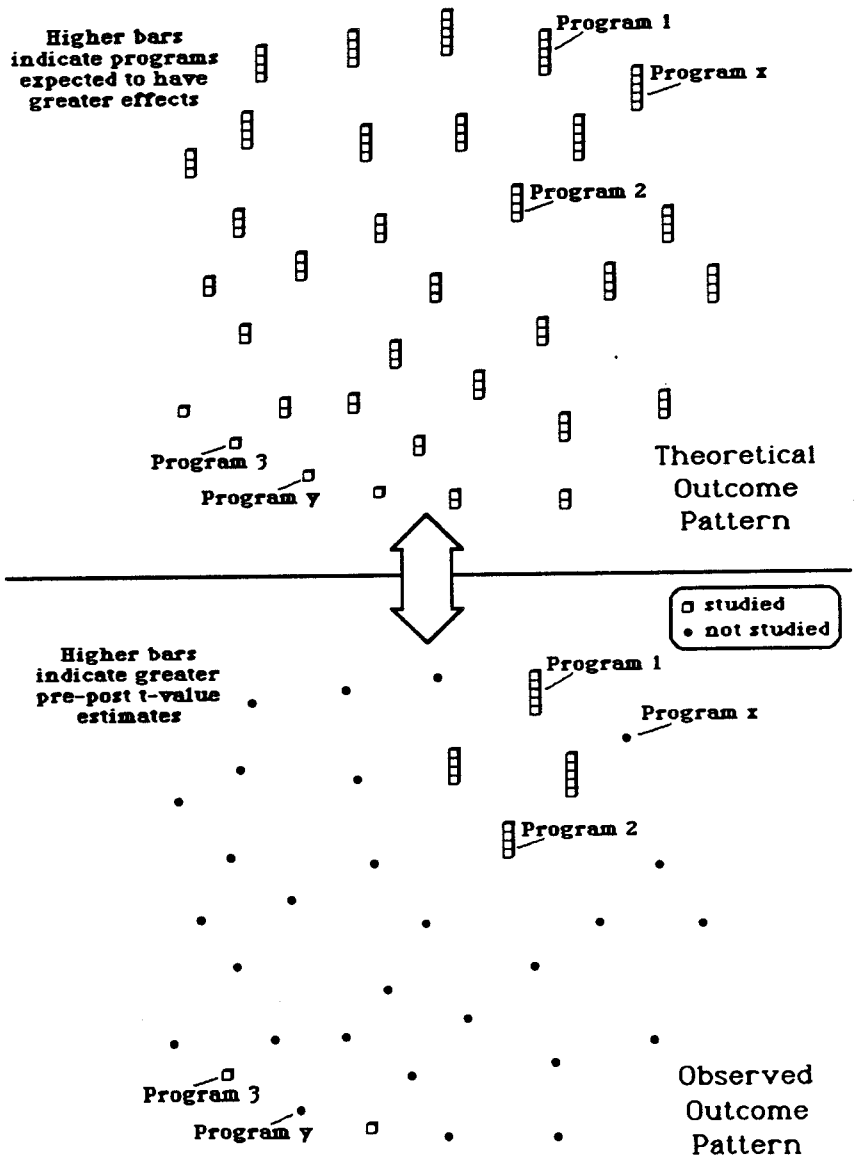

Figure 5. Hypothetical outcome pattern match using a program object pattern.

or geometry ones. We can reach this conclusion only because we have already developed a theory or construct pattern which, through the measurement PM, has confirmed that the nonincluded arithmetic questions are more similar to the included ones than to the spelling or geometry ones.

A second type of outcome PM can be constructed on the basis of a program PM. For instance, if we consider a hypothetical study of three programs or program variations (conceptualized within a broader context of many program types), it would be possible for us to map program effects for any one measure onto the three programs shown in the figure as indicated in Figure 5. The top part of the figure shows the expected outcomes across programs. This map might be generated by performing multidimensional scaling of judges' estimates of the degree of similarity between all potential program variations. The coordinates could then be graphed and average estimates of expected program effect for each program type could be overlayed onto this graph as shown in the top part of the figure. Programs are located on the map on the basis of their shared characteristics. For instance, Programs 1 and 2 might emphasize primarily arithmetic while Program 3 focuses on some other skill (e.g., spelling). Our expectation would be that Program 1 should show the highest effect on measures most related to the nature of that program (e.g., the arithmetic program). Program 3 would be expected to show little or no effect on arithmetic. Furthermore we would expect that the results for Program 2

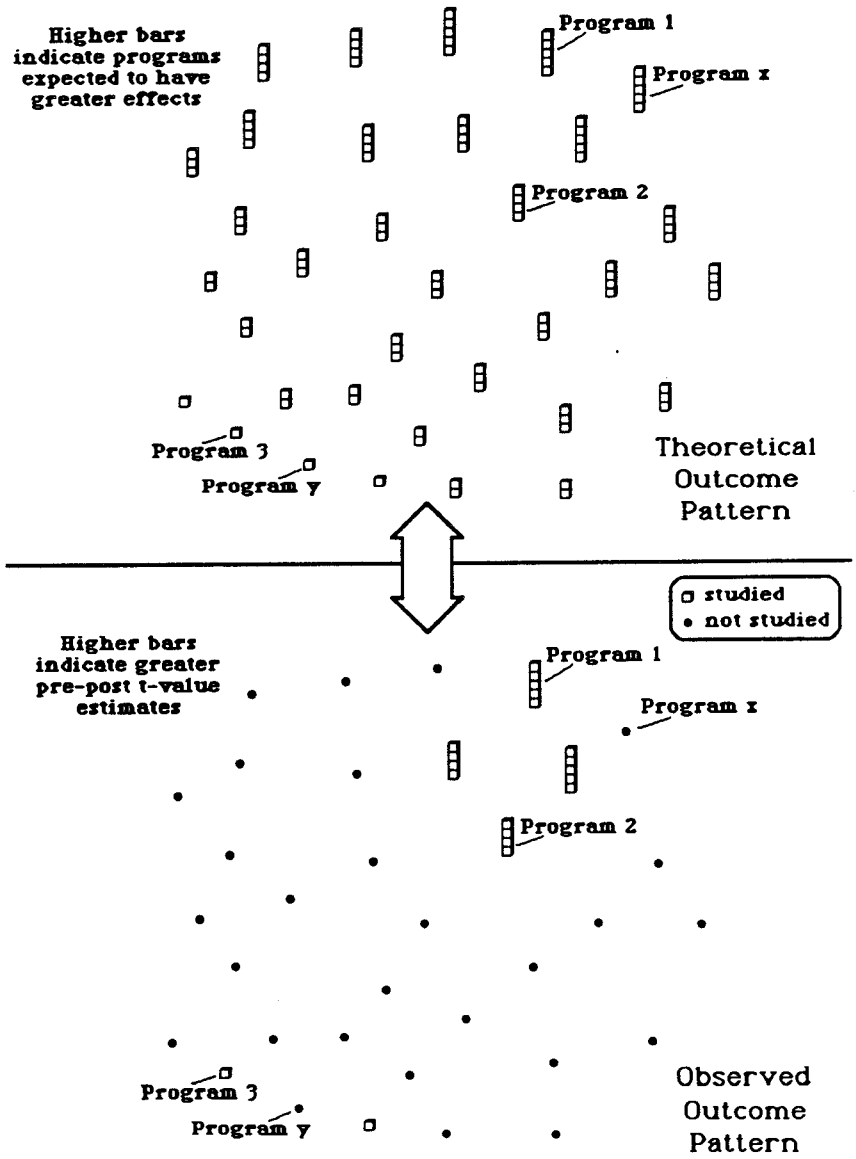

Figure 5. Hypothetical outcome pattern match using a program object pattern.

should be more like those of Program 1 than those of Program 3 because we have already established that Programs 1 and 2 are more like each other.

As in the case of measurement described above we can use this approach to explore the generalizability of program effect across theoretically different types of programs which were not explicitly included in the study. Thus, for a program which we judged as similar to (close to) Program 1 (e.g., labelled Program x on the Figure 5), we would expect the results to be more like those for Program 1. Similarly, for a program which is more similar in nature to Program 3 (e.g., labelled Program y on the figure) we would expect little effect on measures of arithmetic.

Finally, we can also conduct an outcome PM relative to the participants. Again, we assume that we have already developed a pattern of the theoretical interrelationships between participants such that participants who are similar are located closer together on a map than those who differ more. In this case, we would have to state our expectations regarding which types of people are likely to be more affected by the program in question. If we examine gain scores for each participant for a relevant measure (e.g., arithmetic) we may see that the effect is stronger in certain sectors of the map than in others as shown in Figure 6. The top part of the figure shows a mapping of a number of *potential* participants based upon the degree of their shared characteristics. Our theoretical expectations are that certain types of people will be more affected by the program than others, as shown in the top part of the figure. In the bottom of the figure we see that, as expected, the gains are high for persons 1 and 2 and relatively low for person 3. This would be indicative of theoretically interpretable individual differences in program effect. Again, we can use the implicit theory about the similarity between participants and our participant outcome expectations to examine the generalizability of our results across even persons who were *not* in our sample (shown as dots in the bottom part of the figure). Consider in Figure 6 the two persons indicated as x and y neither of whom participated in the study. We would in this case more reasonably conclude that person x *would have been more affected by the program* than person y. We can reach this conclusion only because we have already formulated a theory or pattern which demonstrated that person x is more like the people who were affected by the program than is person y.

When using pattern matching for generalizing, we are assuming, as Campbell (1986, p. 74) does that "nature is 'sticky,' 'viscous,' proximally autocorrelated in space, time, and probably *n*-dimensional attribute space, with adjacent points more similar (as a rule) than nonadjacent ones." Consequently, we would generalize "with most confidence where treatment, setting, population, desired outcome, and year are closest in some overall way to the original program treatment" (Campbell, 1986, p. 75). These assumptions also underlie generalizability as conceived by Judd and Kenny (1981) who state:

> We might conceive of different theoretical populations as departing from the sampled population along a gradient of similarity. Some theoretical populations are quite similar to the sampled population; others are less similar. Generally the confidence that we have in generalization to populations not operationalized depends on the population's location on this gradient of similarity. Such gradients of similarity can also be used for generalization to outcome, treatment and setting constructs that were not operationalized in the research (p. 41).

Thus, theoretical patterns provide a description of the similarity context within which generalizability statements are made.

When we speak about outcomes in a pattern matching framework, we are not speaking about a single outcome or effect, such as a difference between groups on a single measure. Instead, we are looking at the pattern of outcomes based on our theories of the nature of program types, measurement constructs, or participant types some of which are explicitly sampled in our study and some of which are not. The traditional distinction between internal and external validity can be seen to break down here. The generalizability context is used to detect the pattern of outcomes and the outcome pattern is used as the basis for reasonable inferences about generalizability. This occurs because the idea of pattern matching simultaneously encompasses the figure and the ground, the object of focus, and its context.

CONCLUSIONS

We can make some general statements about the idea of pattern matching on the basis of the above discussion and examples. These general statements are implications of a pattern matching approach and can be written as recommendations to researchers on how to improve research through pattern matching. They also serve to provide more concrete statements about what types of theories we need to develop for a research project and how those theories can be utilized in accomplishing the research. Each of these recommendations is preceded by a term which summarizes the nature of the principle involved.

Contextualism: We should articulate and observe our object of interest (i.e., program, measure, participants) within a context. The context should include a range of elements which are more to less related to the object of

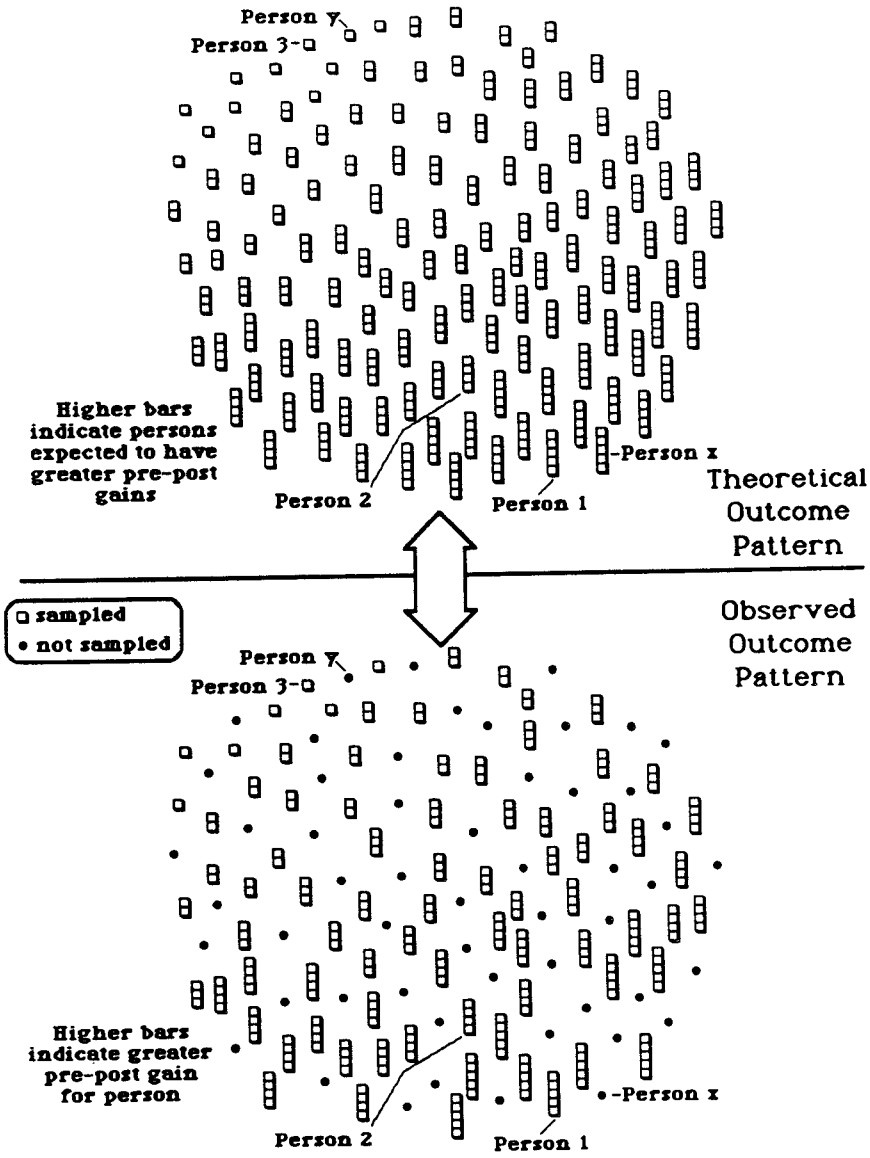

Figure 6. Hypothetical outcome pattern match using a participant object pattern.

interest and, on the theoretical side, should contain objects which are both included in and excluded from the study itself.

Multiplism: We should articulate and observe multiple manifestations of the objects of interest. Pattern matching assumes that wherever possible, multiple programs, measures, and participant groups will be included in a study. This is essentially the same point that is being made in recent work by Cook (1985), Shadish (1986), and Shadish, Cook, and Houts (1986) in their elaboration of a theory of critical multiplism.

Relationalism: We should articulate and observe the relationships between these multiple manifestations of the objects of interest. At the theoretical level, we should attempt to specify how programs, measures, and participant groups are interrelated, including instances of these which are not a part of the study itself. On the observational side we should confirm our ideas about interrelationships by examining the degree of similarity between programs, persons, or measures and matching these to our expectations.

Gradualism: We should articulate and observe the degree of relationship between these multiple manifestations of the objects of interest. This statement is included to encourage us to move away from a categorical view of objects and towards one which views objects as differing in degree on multiple dimensions. This is essentially an encouragement of the "gradient of similarity" idea implied in Judd and Kenny (1981). For instance, instead of arguing that we have three measures

of a single underlying trait, we should recognize that if these are different measures, they by definition vary in some way in measuring that trait. By moving away from a factor-oriented approach we are able to articulate more specific pattern expectations and thus reduce the likelihood that alternative pattern explanations will exist for a given pattern match.

Dualism: We should develop patterns at both the theoretical and observational levels. Pattern matching is based on the linkage between theory and observation. A poor implementation of either of these will lead to poor pattern matching. For all of the above recommendations, the phrase "articulate and observe" was used to emphasize this dualism between theory and observation and stress the importance of activities that must be carried out at both levels.

Parallelism: We should assess the link or match between the theoretical and observed patterns. In order to conduct pattern matching we need two patterns which are parallel in structure. This implies that operationalizations in a research project should be directly based upon a theoretical structure so that we know in advance which part of the theoretical pattern has its corresponding manifestation in any given observation.

Degree of Correspondence: We should examine the degree of correspondence between patterns. A pattern match is never likely to be exact in practice. We need to develop statistical methods which assess the degree to which a match is present. In addition, we need to develop procedures which can be used for diagnosing an imperfect pattern matching. Partial pattern matches will provide a rich source for theory modification and the improvement of operationalizations of the program, measures, and sample.

Pattern matching as described here is still in the early stages of its development although it is based on a long tradition of pattern-based research methodology. As such, it acts as a guide for the development of theory by suggesting how theories might be utilized in the research process. Before pattern matching can reach its full potential as a governing principle in the research process there needs to be additional work on how theoretical patterns might best be constructed, how different types of data can be linked into patterns, and how we might better statistically assess the degree of a match. Furthermore, we must not restrict pattern matching to only the causal hypothesis testing case. All research involves some pattern matching. Sample survey design could be improved by considering measurement and participant pattern matching. Qualitative research could usefully utilize pattern matching as a rubric for categorizing data. Ultimately, our advocacy of the greater use of theory will not progress unless we continue to explore specific strategies like pattern matching which can help guide the development and use of theory in research.

REFERENCES

BICKMAN, L. (Ed.). (1986). Using program theory in evaluation. *New Directions for Program Evaluation*. San Francisco, CA: Jossey-Bass.

CAMPBELL, D.T. (1966). Pattern matching as an essential in distal knowing. In K.R. Hammond (Ed.), *The Psychology of Egon Brunswick*. New York: Holt, Rinehart and Winston.

CAMPBELL, D.T. (1986). Relabeling internal and external validity for applied social scientists. In W. Trochim. (Ed.). Advances in quasi-experimental design and analysis. *New Directions in Program Evaluation*. San Francisco, CA: Jossey-Bass.

CAMPBELL, D.T., & FISKE, D.W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*, 81-105.

CAMPBELL, D.T., & STANLEY, J.C., (1963). Experimental and quasi-experimental designs for research on teaching. In N.L. Gage (Ed.). *Handbook of research on teaching*. Chicago: Rand McNally.

CAMPBELL, D.T., & STANLEY, J.C. (1966). *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally.

CARACELLI, V. (1989). Structured conceptualization: A framework for interpreting evaluation results In W. Trochim (Ed.). *Concept mapping for evaluation and planning*. A Special Issue of *Evaluation and Program Planning, 12*, (1), 45-52.

CHEN, H.T., & ROSSI, P.H. (1983). Evaluating with sense: The theory-driven approach. *Evaluation Review, 7*, 283-302.

CHEN, H.T., & ROSSI, P.H. (1987). The theory-driven approach to validity. *Evaluation and Program Planning, 10*, 95-103.

COOK, T.D. (1985). Postpositivist critical multiplism. In L. Shotland and M.M. Mark (Eds.). *Social science and social policy*. Beverly Hills, CA: Sage.

COOK, T.D., & CAMPBELL, D.T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Chicago: Rand McNally.

CORDRAY, D.S. (1986). Quasi-experimental analysis: A mixture of methods and judgment. In W. Trochim. (Ed.). Advances in quasi-experimental design and analysis. *New directions in program evaluation*. San Francisco, CA: Jossey-Bass.

CRONBACH, L.J. (1982). *Designing evaluations of educational and social programs*. San Francisco, CA: Jossey-Bass.

CRONBACH, L.J., AMBRON, S.R., DORNBUSH, S.M., HESS, R.D., HORNICK, R.C., PHILLIPS, D.C., WALKER, D.F., & WEINER, S.S. (1980). *Toward reform of program evaluation*. San Francisco, CA: Jossey-Bass.

CRONBACH, L.J., & MEEHL, P.E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52*, (4), 281-302.

DAVIS, J. (1989). Construct validity in measurement: A pattern matching approach. In W. Trochim (Ed.). *Concept mapping for evaluation and planning*. A Special Issue of *Evaluation and Program Planning, 12*, (1), 31-36.

DUMONT, J. (1989). Validity of multidimensional scaling in the context of structured conceptualization. In W. Trochim (Ed.). *Concept mapping for evaluation and planning*. A Special Issue of *Evaluation and Program Planning, 12*, (1), 81-86.

EINHORN H.J., & HOGARTH, R.M. (1986). Judging probable cause. *Psychological Bulletin, 99*, (1), 3-19.

GALVIN, P.F. (1989). Concept mapping for planning and evaluation of a Big Brother/Big Sister program. In W. Trochim (Ed.). *Concept mapping for evaluation and planning*. A Special Issue of *Evaluation and Program Planning, 12*, (1), 53-58.

JUDD, C.M., & KENNY, D.A. (1981). *Estimating the effects of social interventions*. Cambridge, MA: Cambridge University Press.

KEITH, D. (1989). Refining concept maps: Methodological issues and an example. In W. Trochim (Ed.). *Concept mapping for evaluation and planning*. A Special Issue of *Evaluation and Program Planning, 12*, (1), 75-80.

MARK, M.M. (1986). Validity typologies and the logic and practice of quasi-experimentation. In W. Trochim. (Ed.). *Advances in quasi-experimental design and analysis. New directions in program evaluation*. San Francisco, CA: Jossey-Bass.

MARQUART, J.M. (1988). *A pattern matching approach to link program theory and evaluation data: The case of employer-sponsored child care*. Unpublished doctoral dissertation, Cornell University, Ithaca, New York.

MARQUART, J.M. (1989). A pattern matching approach to assess the construct validity of an evaluation instrument. In W. Trochim (Ed.). *Concept mapping for evaluation and planning*. A Special Issue of *Evaluation and Program Planning, 12*, (1), 37-44.

SHADISH, W.R. (1986). Planned critical multiplism: Some elaborations. *Behavioral Assessment, 8*, 75-103.

SHADISH, W.R., COOK, T.D., & HOUTS, A.C. (1986). Quasi-experimentation in a critical multiplist mode. In W. Trochim. (Ed.). *Advances in quasi-experimental design and analysis. New directions in program evaluation*. San Francisco, CA: Jossey-Bass.

TROCHIM, W. (1985). Pattern matching, validity, and conceptualization in program evaluation. *Evaluation Review, 9*, (5), 575-604.

TROCHIM, W. (1989). (Ed.). *Concept mapping for evaluation and planning*. A Special Issue of *Evaluation and Program Planning, 12*, (1).

TROCHIM, W., & LINTON, R. (1986). Conceptualization for evaluation and planning. *Evaluation and Program Planning, 9*, 289-308.