

Generating and Generalizing Knowledge about Psychotherapy from Pragmatic Case Studies

TRACY D. EELLS^{a,b}

^a University of Louisville

^b Correspondence concerning this article should be addressed to Tracy D. Eells, University of Louisville,
Department of Psychiatry and Behavioral Sciences, 500 South Preston Street Room 210, Louisville, KY 40202
Email: eells@louisville.edu

*Note: This article is based on a paper presented at the Society for Psychotherapy Research, Edinburgh, Scotland,
June 21-24, 2006 on a panel, "A Marriage Made in Heaven: Systematic Case Studies and Psychotherapy Rresearch,"
Daniel Fishman, Moderator.

ABSTRACT

An epistemological case is made for how single subject psychotherapy research provides unique and untapped opportunities for generating and generalizing scientific knowledge about psychotherapy. The epistemological claim asserts that it is essential that problem and method be aligned in psychotherapy research. Examples of misalignment in experimental and correlational contexts are given and their consequences discussed. Both Molenaar and Valsiner's (2005) genetic metaphor of phenotypes and genotypes and Lewin's (1931) concept of Aristotelian and Galileian thinking provides further epistemological grounds for the value of single subject research. Several suggestions are made for how to reconcile the epistemological problems discussed. Finally, examples are given of how a database generated by the *PCSP* journal process might serve as a tool to generate and generalize psychotherapy knowledge.

Key words: case studies; single-subject versus group psychotherapy research; case formulation; Aristotelian versus Galileian thinking; epistemology

Single subject research can provide a basis for the generation and generalization of scientific knowledge about psychotherapy. In this paper I present the logic of single subject research in building a knowledge base in psychotherapy, followed by a discussion and examples of how a PCSP database might facilitate the aggregation of knowledge about psychotherapy processes and outcomes.

DEFINITION OF SINGLE SUBJECT RESEARCH

Single subject research (SSR) in the study of psychotherapy is *investigation that takes the individual as both the unit of interest and the unit of analysis*. "Unit of interest" means the research question concerns the functioning of an individual entity. The "individual" could be a psychotherapy client, a client-therapist dyad, a therapy group, or an interaction among these.

Not all questions that psychotherapy researchers ask are best answered or even can be answered through SSR. For example, the questions, “Is psychotherapy effective?”, “Is treatment A more effective than treatment B?”, and “What is the contribution of component A in treatment B to outcome?” are best addressed from an inter-individual or socio-ecological frame of reference (Valsiner, 1986b). They are most frequently addressed by randomly assigning individuals to treatments, delivering those treatments, and comparing differences between treatments using analysis of variance in the context of null hypothesis significance testing. They might also be addressed through meta-analysis, which bypasses significance testing, instead looking at the magnitude of differences between groups. As typically investigated, each of these questions is focused on one or more groups of individuals, each represented by a statistical “average” client, and the analytical tools are commensurate in that they are inter-individual in nature, which is to say that they fundamentally produce information comparing an average person to another. Policy makers and leaders of mental care delivery systems are most likely to be interested in questions of this nature since they make decision affecting populations of individuals, such as those related to health care delivery. For example, an answer to the question, “Do interventions that are tailored to an individual’s readiness to change affect outcome?”, may affect how policy makers allocate tax dollars to assist the mentally ill.

In contrast, other questions are uniquely suited to SSR. For example, “What accounts for production and maintenance of A’s presenting problems?”, “How might therapist B best help person A?”, “What is the best intervention for person A in context X?”, “What course did person A follow in his/her recovery or failure to recover?”, and “What intentions drive A’s behavior?” Note these questions are embedded in specific therapeutic contexts and refer to specific, unique individuals. Note further that the practice of psychotherapy always occurs at the level of the individual, and it is questions at this level that are of fundamental concern to practicing psychotherapists. Although questions framed in this manner may make psychologists trained in traditional research methods and statistics courses skeptical about the generalizability and replicability of any answer to other persons or contexts, one goal of this paper is to make the case that such answers can be generalized.

The term “unit of analysis” refers to the level of analysis that the research methodology addresses and that limits the scope of inference. For example, as noted with regard to the first set of questions, the group is the unit of analysis in analysis of variance, since group means are compared. The group is also the unit of analysis with correlational research since a correlation measures one’s standing in a group of scores relative to another distribution. These are fundamentally methodologies developed for an inter-individual frame of reference (Valsiner, 1986b). According to Valsiner (p. 396), the inter-individual frame involves a comparison of an individual subject (or samples of subjects) with other individuals (samples) in order to determine the standing of these subjects relative to one another. Statements such as, “The cognitive-behavioral group responded better than the psychodynamic group” or “The average person receiving psychotherapy is less symptomatic 16 weeks later than 75% of those not receiving treatment” reflect the inter-individual frame of reference. These conclusions provide comparative information, but do not address intra-individual issues such as how variable X

interacts with variables Y or Z within an individual; nor do they address questions about client A's motivations, preferred problem-solving strategies, self concept, concept of others, or style of managing stress. Similarly, the inter-individual context does not address questions about how individuals interact with specific sociocultural environments, including behavior in a specific client-therapist dyad. Although useful for answering questions about differences *between* systems, the inter-individual frame does not address variation *within* the systems that are compared, except as error variance. In both ANOVA and correlational analysis, the scope of inference is limited to the unit of the group. As noted, in the case of ANOVA, it is limited to the behavior of the average individual in a group, not the behavior of any specific individual. In correlation analysis, the level of inference is limited to a person's score on a variable relative to another variable, in the context of the group in question. It does not address the degree of intra-individual variation of two variables.

The definition of single subject research does not exclude multiple single subject studies, either sequentially or concurrently. For example, an experimenter conducting a study with 20 individuals, analyzing each individual's data on an intra-individual basis, then aggregating the findings of these 20 separate analyses, would satisfy the definition of SSR given in this paper.

ALIGNING PROBLEM AND METHOD

The epistemological grounds for asserting that SSR can provide a basis for generating and generalizing psychological knowledge begins with the claim that the problems we wish to solve and the means we use for solving them must be aligned. When they are not, which is often the case in psychological research, a mismatch between problem and method is created. Ludwig Wittgenstein (1953) expressed this idea in his *Philosophical Investigations*: "In psychology there are experimental methods and *conceptual confusion*. The existence of the experimental method makes us think we have the means of solving the problems which trouble us; though problem and method pass one another by" (p. 232). Wittgenstein's point has been restated and demonstrated in various forms and contexts for more than 50 years (e.g., Bakan, 1967; Sidman, 1952; Thorngate, 1986; Valsiner, 1986a). Valsiner (1986a), for example, empirically demonstrated that people with varying levels of psychological and statistical training commonly interpret correlational data in terms of individual psychological mechanisms.

Sidman (1952) demonstrated mathematically that the shape of individual learning curves, specifically those based on Hull's exponential growth function, cannot be described by nor derived from group data. Inter-individual variability in learning curve asymptotes and in the rates of approach to these asymptotes preclude such inferences from group data. He further claimed that the same conclusion can be reached about other functional relations, a point later demonstrated by Bakan (1954). This point is relevant today since functional analysis is a key analytical tool used by cognitive-behavioral therapists to understand their clients (Koerner, 2007; Nezu, Nezu, & Lombardo, 2004; Persons & Tompkins, 2007). Noting the work of Sidman and Bakan, Thorngate (1986) concludes that "patterns of aggregates have no necessary connection

with aggregates of patterns” (p. 73) and that “what occurs on average is not necessarily what occurs in general” (p. 75).

To further elucidate this issue, Bakan (1967) distinguished between general-type propositions and aggregate-type propositions. The former “asserts something presumably true of each and every member of a designable class” (p. 35). Examples might be, “A variable ratio schedule of reinforcement will lead to high levels of activity”, “A fixed interval schedule of reinforcement will lead to accelerated levels of activity as the interval comes to an end”, “Biological predisposition, early life trauma and an invalidating interpersonal environment characterized by intense negative affect will lead to borderline behavior”, and “After a choice has been made between two closely desired outcomes, preference for the chosen outcome will increase and preference for the outcome not chosen will decrease.” An aggregate-type proposition “asserts something presumably true of the class considered as an aggregate” (p. 35). Examples include “80% of those completing cognitive-behavioral treatment for panic disorder no longer experience panic attacks”, “32% to 44% of those completing behavioral or cognitive-behavioral treatment for smoking cessation maintain abstinence for at least 12 months”, and “Client characteristics are the strongest predictor of outcome in psychotherapy, followed by the therapeutic alliance.” Note that aggregate-type propositions are probabilistic in nature, whereas general-type propositions are framed in causal terms. Bakan noted further that the distinction between the two types of propositions is sharply illustrated by the meaning of the “next” case. For general-type propositions, the “next” case always poses a threat to falsification of the proposition, just as Popper (1959) observed that the proposition “All swans are white” is disproved by the observation of a single black swan. If a general-type proposition fails to be confirmed by a member of the class to which the proposition presumably applies, the proposition must either be revised, rejected, or the class further narrowed. Bakan noted as well that statistical propositions are almost always aggregate in type, and that statistical inference is not a way of making inferences from one type of proposition to the other. Rather it permits one to make inferences to a population aggregate on the basis of a sample aggregate.

The assumption underlying the practice of stating hypotheses in terms of individual processes but testing these hypotheses from an inter-individual framework seems to be that in understanding inter-individual differences we learn something about intra-individual processes and individual-environment processes as they unfold through time. The practical consequences of misaligning problem and method in psychological research are so significant, however, that a fully developed empirical demonstration of the problem is worthwhile. I will demonstrate these consequences first in the experimental context, then in the correlational context.

The Experimental Context: The Example of Causal Attribution

Causal attribution has long been an area of research in social psychology. The basic question is, “What determines an individual’s attributions of causes to actions?” Kelley (1967; 1973) suggested that three types of information function independently in this regard: consensus, distinctiveness, and consistency information. When consensus, distinctiveness, and consistency information are all high, for example, the theory holds that we attribute causality to

the stimulus. If I say “John laughed at the comedian” and you know that others in the audience also laughed (consensus), that John tends not to laugh at other comedians (distinctiveness), and that John has usually laughed at this comedian (consistency), you will likely attribute his behavior to the comedian.

McArthur (1972) tested Kelley’s theory by randomly assigning individuals to one of eight groups. She described a behavioral event to each group, including information about consensus, distinctiveness and consistency, then asked each person what probably caused the event to occur. Each group was provided information so that each of the eight possible combinations of the three independent variables was represented by a group. Using analysis of variance techniques from the inter-individual frame of reference, McArthur found general support for Kelley’s model. She concluded that the study has “clearly contributed to our understanding of peoples’ beliefs about the causes of another person’s behavior: *how* they decide *why* someone behaved as he did...” (p. 193). Although the words, “peoples’ beliefs” may suggest that group psychological processes are the focus of the paper, the implication is clear that the findings are relevant to how individuals attribute causality to behavior. The key question remains: do these effects generalize to individual psychological functioning?

To answer this question, Tukey and Borgida (1983) attempted to replicate the findings from the intra-individual frame of reference. Rather than giving partial information to each of eight groups, they gave all the eight combinations of information to each of 25 subjects, and then conducted 25 separate intra-individual analyses. Assuming that Kelley’s theory is correct about how we attribute causality, one would expect findings similar to those of McArthur. Instead, Tukey and Borgida identified eight primary attribution patterns among the subjects compared to the one pattern in the McArthur study. The responses of only three out of the 25 subjects in the intra-individual analysis fit the best fitting general model in the inter-individual analysis, certainly not a pattern that can be described as general.

What explains these differences in outcome? One possibility is sampling error, since the two studies were done with different subjects. However, Tukey and Borgida also analyzed their data on an inter-individual basis in order to determine how the typical or average subject responded. The resulting pattern was identical to the main effects observed in the McArthur study, which as noted characterized only three individuals based on results of intra-individual analysis. A second explanation might be that chance variation resulted in the eight patterns identified in the intra-individual analyses. The strategy was to analyze each individual subject, then aggregate the findings. Maybe these subjects are not representative and thus one cannot generalize to others. Perhaps one cannot make statements about causality because, unlike the inter-individual study, no random assignment of subjects to groups occurred. Maybe their responses are not reliable. This explanation is not satisfactory, however, when one considers that over 400 data points were obtained from each individual, that chance effects were taken into account in the analysis, and that each individual responded to items under multiple conditions, thus serving as his or her own control.

Another explanation for the difference between the two types of analyses is to view the latter in an individual differences framework. Individual differences research has a longstanding tradition in psychology. Cronbach (1957) identified it as one of the "two disciplines of scientific psychology" nearly 50 years ago and it remains so today. It is true that individual differences exist, but this explanation is more a dismissal of the discrepancies between these two types of studies than a true explanation. First, the conclusions to the inter-individual study are stated in general terms, but clearly are not general findings. Second, the intra-individual study takes an approach to individual differences research that is different from the usual approach. More typically, individual differences research is done in a correlational paradigm where single observations are gathered from each individual on each variable and these observations are correlated. That is, most individual differences research is done in a framework in which data are aggregated first and then analyzed. In the Tukey and Borgida study, individual differences are revealed through intra-individual analyses which are then aggregated to indicate individual differences.

As Wittgenstein asserted, it does appear that in an experimental or quasi-experimental context, problem and method do, on occasion, "pass one another by." Next, we will examine whether the same is true in the correlational context.

***The Correlational Context:
The Example of Testosterone Level and Orgasmic Activity***

Suppose you are interested in the relationship between testosterone level in blood plasma and orgasmic activity. Consider three approaches to examining this relationship. First, we draw blood from a sample of young men and determine their testosterone level. A day later we contact each subject and ask whether he has had an orgasm within a period beginning 32 hours prior to the blood drawing and ending 16 hours after. If he has, we record a "yes"; otherwise, we record a "no". A point-biserial correlation coefficient can tell us whether there is a relationship. When Kraemer (1978) conducted this study with a sample of 20 men, she found a correlation coefficient of .08, suggesting no relationship between testosterone level and orgasmic activity.

Recognizing that testosterone level has substantial intra-individual variability, one might approach the question in a different manner. Instead of drawing just one sample, we draw several over the course of two months. The young men come in every other day to have their blood drawn. We also determine, as before, whether they had an orgasm within a period beginning 32 hours prior to having the blood drawn and ending 16 hours after. For each individual, we then compute averages for testosterone level and the percentage of "yes" responses over the two month period. When these two averages were associated across subjects, the Pearson correlation was $-.44$ ($p < .10$). The Spearman rank correlation was $-.56$ ($p < .05$). Either of these suggests a negative relationship between testosterone level and orgasmic activity. Note that both of these correlations are inter-individual. As is usually the case in correlational research, the correlation coefficient is calculated by taking a single observation from each subject on each of the two variables.

Contrast these approaches with a third: calculate a point-biserial correlation coefficient for each subject. One can do this because Kraemer gathered several measures of both testosterone level and orgasmic activity from each participant. (One participant had to be dropped from this part of the analysis because he reported only one orgasm during the two months of the study.) Of the remaining 19 participants, 16 had positive correlations. A sign test was significant at $p < .01$. The median increase associated with orgasmic activity expressed as a percentage of the intra-individual standard deviation is 38%. This indicates a *positive* relationship between testosterone level and orgasmic activity.

What should one conclude from these data? The answer depends on the problem one seeks to solve. If one is interested in the direct relationship between two intra-individual variables, such as hormonal activity and changes in libido, the method should be directed at the intra-individual level. Questions related to changes in the relationship between the interaction of two intra-individual variables as they differentiate individuals from each other should be addressed by an inter-individual correlation.

PSYCHOLOGICAL PHENOTYPES AND GENOTYPES

Returning to the central question of this article, the task of generating generalized knowledge about psychotherapy on the basis of single subject research is, of course, complicated by the nature of the object of study – the human being interacting in a physical and social context. Molenaar and Valsiner's (2005) illustrate this well. They aptly describe each individual human as a temporally organized system of varied complexity that is unique in all aspects. Further, each person follows an idiosyncratic path through life, a path which is both crucial to that individual's development, but also entirely individual and unlikely ever to be repeated again.

Given the idiographic nature of each person and his or her unique life path, how does one proceed with an effort to generalize? Molenaar and Valsiner answer this question using the genetic metaphor of a phenotype and genotype. As they describe it, generic, universal processes co-occur with the unique features of individuals. These processes are evolutionary in nature; they make uniqueness possible since uniqueness is a product of adaptation and adaptation is crucial to an individual's survival. By their account, the generalization task in psychology, therefore, is to identify these generic, or genotypic, processes by studying their case-specific, or phenotypic, representations, that is, human beings.

Further appreciation of Molenaar and Valsiner's genetic metaphor and the "embeddedness" of the individual in his or her interpersonal, cultural, and physical environment, and the challenge this presents to generating generalized scientific knowledge is gained from examining the work of Kurt Lewin. We undertake this task in the following section, then review research approaches appropriate to the task.

PSYCHOLOGY OF THE CONCRETE SITUATION

Kurt Lewin (1931) was one of the first psychologists to advocate the single subject design as a method of choice. His call for single or “pure” case studies, was embedded in his broader appeal for a “Galileian” rather than an “Aristotelian” psychology.¹ (Note that all the quotes in this section are from Lewin, 1931.) One must have a general understanding of Lewin’s concepts of Galileian and Aristotelian thinking in order to understand the basis of his claim that case studies are necessary to generate generalized knowledge in psychology. Thus, I will review these concepts.

Aristotelian Thinking

Lewin (1931) claimed that the contemporary psychology of his time was “thoroughly Aristotelian” (p. 151) and to advance, should become more Galileian. By Aristotelian, he meant the following: (1) that class and essence are equated; (2) that the criterion and expression of lawfulness is frequency, that is, frequency is taken as the overriding criterion to establish lawfulness; (3) that “historic-geographic” concepts (i.e., concepts associated with frequency in certain times and places) are utilized, and consequently, significant single case results are viewed either as fortuitous or anomalous, but either way, are typically ignored.

Class and Essence

Conflating class and essence -- the essential nature of a thing -- is a fundamental Aristotelian error, according to Lewin. Aristotle defined class as “the sum total of those characteristics which a group of objects have in common” (p. 144). The essence of gold might be defined as an element having an atomic number of 79. All elements having this atomic number are included in the class and all those with different atomic number are excluded. Thus, “atomic number of 79” defines both the class and the essence of gold. The fallacy in this mode of thinking, according to Lewin, only becomes apparent when the objects that make up a class are not homogeneous like gold, but instead are heterogeneous like people. Lewin provides an example from developmental psychology:

Whatever is common to children of a given age is set up as the fundamental character of that age. The fact that three-year-old children are quite often negative is considered evidence that negativism is inherent in the nature of three-year-olds, and the concept of a negativistic

¹ Lewin used “Aristotelian” and “Galileian” as umbrella terms to communicate a broad set of ideas about philosophy of science, scientific lawfulness, and research strategy. His concern was the modes of doing science that grew from Aristotle and Galileo, not to their specific writings. Further, Lewin explicitly rejected the wholesale importation of methodologies from physics.

age or stage is then regarded as an explanation (though perhaps not a complete one) for the appearance of negativism in a given particular case! (p. 153)

Similar examples can be given from trait theory and statistics. It is an error, for example, to explain an individual's aggressiveness by attributing causal status to an aggression trait. In statistics, mean results may be conflated with the scores that comprise the mean, and the mean is interpreted as conveying a “deeper”, more significant, or more essential quality of a group than the scores themselves. The central limit theorem in statistics illustrates how false this is. A sampling distribution of the mean, as sample size grows, always produces a normal curve regardless of how non-normal the distribution of the population scores may be.²

Aristotelian Lawfulness

Frequency is the cornerstone of Aristotelian lawfulness, according to Lewin. His example of three-year-old negativism illustrates that the Aristotelian concept of classification is closely connected with that of lawfulness. It is the frequency of observing negativism in three-year olds that led to the “law” of inherent negativism in children of this age. Lewin asserts that this “extravagant overvaluation of repetition” in which frequency is viewed as the “criterion and expression of lawfulness” (p. 152-153) is fundamental to an Aristotelian mode of thought: The discovery of regularities is the sine qua non of laws in psychology. By virtue of this, the single case is typically discounted. It may represent an event that does not occur frequently, and thus considered outside of scientific purview. Even if the case does represent a frequently recurring event, there is no telling the extent of its frequency from the single case alone. (Note that “generality” is viewed as generality to different individuals, not generality to similar processes in different individuals or similar processes in the same individual, but at a different period of time.)

Lewin further asserts that the use of statistics in psychology is the most striking expression of the Aristotelian mode of thought. This is directly related to the notion of lawfulness expressed through frequency because statistical methods rely upon multiple observations. According to Lewin, the fundamental explanation for the reliance on frequency in psychology is skepticism that psychology is a fully lawful discipline. Thus, by default, some criteria must be decided upon to determine what will be classified as lawful and what will not. As Lewin writes, “Any psychology that does not recognize lawfulness as inherent in the nature of the psychic, and hence in *all* psychical processes, even those occurring only once, *must* have criteria to decide...whether or not it has in any given case to deal with lawful phenomena” (p. 152). The reliance on frequency as the criterion for lawfulness also has consequences for

² Russell's “theory of logical types” communicates a similar point. In *Principles of Mathematics*, Russell writes that a class and the members comprising it are qualitatively different and that equating the two is an error of logic.

standards of proof in psychology. Aristotelian methods give rise to a laxity in standards of proof, asserts Lewin, because conclusions are always stated in terms of the "average" which permits any single case that goes counter to the "average" trend to be discounted. In effect, it is excluded on ground of being not lawful.

Historical-Geographic Concepts

Lewin also saw a focus on historical-geographic concepts as characteristic of Aristotelian concepts. This relates to the idea of frequency as the basis for lawfulness and on the embeddedness of frequency in certain times and places. Events repeating in time in a regular manner are what interest the Aristotelian psychologist. Conversely, events which occur only once lack temporal and locational ballast and therefore are considered lacking in lawfulness and of little or no significance. In Lewin's words, "Regularity or particularity was understood entirely in *historical* terms... The complete freedom from exceptions, the 'always'.... still has here its original connections with the frequency with which similar cases have occurred in the *actual, historical* course of events in the everyday world" (p. 146). To illustrate, he provides an example,

Light objects, under the conditions of everyday life, relatively frequently go up; heavy objects usually go down. The flame of the fire, at any rate under the conditions known to Aristotle, almost always goes upward. It is these frequency rules, within the limits of the climate, mode of life, etc., familiar to Aristotle, that determine the nature and tendency to be ascribed to each class of objects, and lead in the present instance to the conclusion that flames and light bodies have a tendency upward. (p. 146) . . .

[In psychology,] the historical bent of psychological concepts is again not always immediately obvious as such, but is bound up with non-historic, systematic concepts and undifferentiated from them. (p. 156) . . .

[Further, this] quasi-historical set forms...the central point for the understanding and criticism of this mode of concept formation (p. 157).

Lewin noted that psychological concepts dependent upon historical-geographic dimensions may mislead. An analogy in physics would be to study hydrodynamics not in the laboratory but by investigating the largest rivers in the world. He concludes that the reliance upon historical-geographic concepts lead Aristotelian-minded psychologists to overvalue the historically important and to disregard the "ordinary", and secondly, to over-regard frequency as an index of what to study. The historic become conflated with abstract classes and the "full reality of the concrete case" is neglected.

Aristotelian View of the Single Case

From the Aristotelian standpoint, the scientific value of a single case is discounted on grounds of its lack of frequency. Its lack of scientific standing is further substantiated when class and essence (i.e., features of the members of the class) are not differentiated. From this standpoint, attending to even interesting and exceptional single cases seems "a scientifically

unimportant bit of folly" (p. 151) to the Aristotelian researcher who holds to the belief: "That which does not occur repeatedly lies outside the realm of the comprehensible" (p. 152).

Galileian Thinking

Lewin's call for a Galileian mode of thinking in psychology is fundamentally a call for the analysis and understanding of the "concrete situation." According to Lewin, the task of psychology should be "the providing of a workable representation of a *concrete psychological situation* according to its individual characteristics and its associated functional properties, and of the concrete structure of the psychological *person* and its "internal dynamic facts" (p. 174). The principle underlying this suggestion is that "*the dynamics of the processes are always to be derived from the relation of the concrete individual to the concrete situation*, and, so far as internal forces are concerned, from the mutual relations of the various systems that make up the individual" (p. 174).

The focus on the situation has significant implications for scientific concept formation, for determining lawfulness, and for valuing importance to single subject research. The key implication for scientific concept formation is that the Galileian approach allows for phenotypic and genotypic concepts, whereas the Aristotelian approach does not. For example, consider an ambivalent attachment that a psychotherapy client may feel toward his/her therapist. The client is both drawn toward and away from the therapist. In Freudian terms, this may represent positive and negative transference. In cognitive-behavioral terms, it can be understood as a combination of expected rewards or punishments coming from caretakers or as contradictory, maladaptive cognitions. This genotypic process can manifest itself in a variety of phenotypic ways. For example, one client, motivated by a view of the therapist as an understanding and caring listener, might resolve before a session to discuss his shame about feelings of dependency. During the session, however, he may fear rejection by the therapist and fail to raise this topic. Another client may have similar motivations, but nevertheless raise the topic. In these cases the genotypic processes are similar, but the phenotypic expressions are significantly different. A purely Aristotelian approach might involve enumerating these instances based on their historical-geographic frequency and categorizing them in different classes, say "self-disclosing behaviors" versus "avoiding behaviors."

Galileian Lawfulness

Lewin argued that the Galileian view of lawfulness is much stricter than the Aristotelian. The view is of "strict exceptionless lawfulness" (p. 159), with the prescription that "every psychological law must hold without exception" (p. 159). Exceptions to a law are not regarded as anomalous or due to chance, but represent a potentially fatal threat to the law. Lewin observed that the shift from Aristotelian to Galileian physics gave the latter

its intoxicating feeling of infinite breadth, because it does not, like the abstract class concepts, level out the rich variety of the world and because a single law embraces the whole field. (p. 159)

The Galileian View of Single Subject Research

Lewin advocated a shift in thinking from the average to the “pure” case. The average case is based on frequency and the historical-geographical setting. The pure case, in contrast, is illustrated by how the law of movement on an inclined plan is validated in physics. It is not done by taking the average of real stones rolling down hills and then considering this average to as the most probable case. Rather, it is based on “the ‘frictionless’ rolling of an ‘ideal’ sphere down an ‘absolutely straight’ and hard plane, that is, upon a process that even the laboratory can only approximate, and which is most extremely improbable in daily life” (p. 161). In effect, one disregards repeatedly observable facts, focusing instead on the most pronounced exceptions that best illustrate a general law applicable to every case, but only obscurely so in the “average” case.

According to Lewin, one finds “pure cases” in concrete situations. The person and processes being studied should not be decontextualized from their environment. The situation assumes as much importance as the person; the two are together considered part of an irreducible, inseparable whole. This view contrasts with the Aristotelian view in which the situation is considered as a factor that may modify the object of concern, but does not significantly alter its constituent parts.

An analysis of the concrete situation has implications for psychological lawfulness. General laws of processes must take into account all possible situations. It is important to comprehend “the whole situation involved, with all its characteristics, as precisely as possible” (p. 166). Further, the “general validity of the law and [the] concreteness of the individual case are not antitheses” (p. 175) but rather the properly chosen individual case confirms the law, and repetition becomes superfluous. A further implication of the analysis of the concrete situation is how variability is considered. From the Aristotelian standpoint, which focuses on class and essence, variability is considered a nuisance or irrelevant. In contrast, from the Galileian standpoint, variability is the object of interest since it resides in the concrete situation; it is what characterizes nature and what must be understood in order to arrive at lawfulness.

Lewin sums up his call for a Galileian mode of thought in psychology by pointing toward the Aristotelian approach with its “too meager scientific courage in the question of the *lawfulness* of the psychological, too slight demands upon the validity of psychological laws, and the tendency, which goes hand in hand with this leaning toward mere regularity, to specifically *historic-geographic concepts*” (p. 174).

Reconciling Aristotelian and Galileian Thinking

Is it realistic or even desirable to consider a shift toward a more Galileian mode of thought in psychotherapy research? While this paper is not the venue for a comprehensive examination of this question, a few points emerge as particularly relevant to single subject research. First, it is possible that the Aristotelian and Galileian perspectives might serve as constructively opposed viewpoints that, when taken together mindfully and strategically, could help advance knowledge through their unique perspectives and the particular kind of knowledge each tends to generate.

Second, Lewin was correct in his prediction of the ascendancy and dominance of the Aristotelian framework. Psychological science, like other social sciences, has been overtaken by the “probabilistic revolution” (Kruger, Daston, & Heidelberger, 1987), in which probabilistic statements about lawfulness have replaced deterministic views of psychological lawfulness. The Galileian idea of lawfulness in psychology sets a stricter standard than what predominates today. It also underlies examples showing that a single, well-selected case study can contribute to scientific progress, as illustrated by the law of movement on a inclined plane described earlier. Some psychological knowledge may meet strict Galileian requirements of lawfulness. Examples might include knowledge about basic sensation-perception processes such as dark adaptation, perceptual illusion phenomena, processes in psychophysics such as a just noticeable difference, and illusory movement in visual perception. In the field of learning, which underlies cognitive-behavioral therapy, principles of operant and classical conditioning may also have reached the level of general lawfulness in psychology. In several of these areas, single subject research actually played a central or defining role (Morgan & Morgan, 2001).

Third, in areas where phenotypic-genotypic relationships have not been established, one constructive goal may be to agree on or aim to agree on existing phenotypes that might then be explored for their underlying characteristics. For example, whereas much remains to be discovered about the genotypic processes underlying memory, agreement appears to exist about basic phenotypic memory phenomena, such as primacy and recency effects, retroactive and proactive inhibition, reconstructive phenomena, and common errors of memory such as biasing, transience, absent-mindedness, and tip-of-the-tongue phenomena (Schacter, 2001). Another possible example derives from Newell and Simon’s study of human problem solving (1972), particularly Chase and Simon’s (1973) study of expert performance. While agreement seems to exist among cognitive psychologists that experts outperform non-experts in domain specific tasks, there is no consensus regarding the underlying cognitive processes.

In the field of psychotherapy, researchers might aim for consensus on phenomena that occur in treatment, and then seek to understand them. Examples might include dyadic processes, coping or defensive phenomena, the role of narrative in therapy, or phase-related phenomena such as those occurring at the beginning, middle and end of therapy. Single subject research can play a key role in these processes since it more readily offers the opportunity to explore micro-behavioral sequences. Hints as to these processes may come from group-focused

outcome treatments that might then be explored more closely in case studies. Alternatively, the effect of micro-behavioral processes might then be explored as to their effect on outcome in larger-sample studies.

Finally, single subject research designs can help correct the epistemological problems described in this paper and also help build a more Galileian psychotherapy research base, in order to counterbalance the predominant Aristotelian mode of thinking. Several strategies have been recommended. First, it should be clear that researchers must remain vigilant about the frame of reference in which they are working and what inferences can and cannot properly be made from it. In general, one should not make inferences across frames of references unless one has an evidentiary base for doing so. Second, the strategy of "analyze then aggregate" rather than "aggregate then analyze" might be used more often (Thorngate, 1986). This strategy requires obtaining multiple observations from single individuals, analyzing those data at the level of the individual, then aggregating to support the nomothetic goals of psychology. This is the strategy followed by Tukey and Borgida (1983) and Kraemer (1978), as described earlier. A third strategy is to expand our use of methodological tools beyond the current dependence on statistical inference and null hypothesis testing as the crucible for creating psychotherapeutic knowledge. In light of the controversy surrounding null hypothesis testing (Krueger, 2001; Morrison & Henkel, 1970), it is somewhat surprising that it remains as popular as it does. Case study methodologies can offer a viable alternative that still permits inductive inferences from findings (Flyvbjerg, 2006). Further, as Gigerenger (1991) has documented, researchers have tended to convert their analytical tools into theories used to explain the very phenomena they study. A strategy of methodological pluralism helps guard against this possibility.

PCSP AS A TOOL TO ADVANCE PSYCHOTHERAPY KNOWLEDGE

In light of the perspectives examined above, the *Pragmatic Case Study in Psychotherapy* (PCSP) journal offers a unique opportunity to contribute to our knowledge base about psychotherapy. A powerful feature of the PCSP journal is the database of cases, all presented in a similar organizational format, that is accruing with the publication of each issue. In this section, I will describe some ways that a large database of this type might be used to build psychotherapy knowledge.

Identification of Anomalous Cases

Anomalous cases are those that do not follow the general trend of a set of findings. They may be exceptional either in their success or failure. It is possible that further study of these cases could reveal patterns undetected in the general findings. For example, studies of cognitive-behavioral therapy for individuals with schizophrenia have suggested that a primary effect of successful treatment is a reduction in the conviction in the reality of hallucinations and delusions, rather than the elimination of these phenomena (e.g., Sensky et al., 2000). It is possible that a case study of someone who completely abandoned psychotic thinking (i.e., a "super-responder") might reveal a set of factors that led to the extraordinary outcome. Similar case studies of

exceptionally successful therapists might also reveal elements of successful treatment that may generalize to therapists with more ordinary outcomes. Ricks' (1974) case studies, for example, identified behaviors that distinguished two therapists, one with excellent outcomes and the other with less successful outcomes (see also Okiishi, Lambert, Nielsen, & Ogles, 2003). Case studies of treatment failures, which have been estimated to occur in approximately 8% of those treated (Lambert & Ogles, 2004), might also reveal important processes to avoid. Treatment failure is a particularly difficult phenomenon to investigate using experimental designs. These exceptions and outliers, though sources of "error variability" in one sense, may very well represent important sources of meaningful variability in a broader sense (see also Gould, 1996).

Identification of Serendipitous Explanations

There are many examples of significant serendipitous discoveries in science (Merton, 1989), including psychology, where Skinner (1956), for example, credited serendipity for his discovery of the effect of schedules of reinforcement on behavior patterns. As described by Tang and DeRubeis (1999), recognition of the phenomenon of sudden therapeutic gains by examining individual response rates to cognitive therapy for depression appears to be an example of serendipity in the psychotherapy research literature.

Since it prizes variation as a focus of study rather than "error" to be minimized, single subject research is ideally suited for capitalizing on serendipity. It is possible that a review of a large database of similar cases may reveal patterns that the study of individual cases would not. For example, an influential cognitive model of social phobia proposes that this condition is the product of negative cognitions and images that lead to heightened self-consciousness and inhibited spontaneity that limits effective social interactions (Clark & Wells, 1995). What if a review of several *PCSP* case studies of individuals treated for social phobia with cognitive-behavioral techniques, also suggested similar patterns of disruptions in self concept and maladaptive defensive/coping style among these individuals, as suggested by Wolfe's (2005) integrative model of treatment for anxiety disorders? A broader conceptualization of social phobia may result, leading to more successful outcomes. As another example, a closer examination of individuals exhibiting "sudden gains" in therapy may also yield unexpected insights (Stiles et al., 2003; Tang & DeRubeis, 1999). Of course, predicting a specific serendipitous finding is contrary to the concept itself. Nevertheless, these examples illustrate a kind of pattern detection that becomes possible when a set of case studies, each presented following a similar format and focusing on the same condition, is juxtaposed and examined.

Follow the Maxim of "Just Sufficient Isolation of Complex Processes"

Cronbach (1986) recommended the strategy of isolating complex phenomena sufficiently to allow systematic study of them, but not so much that important relational effects cannot occur or come to the attention of observant researchers. He illustrated the strategy in reference to Roethlisberger and Dickson's (1939) study of worker productivity, which led (serendipitously) to the discovery of the Hawthorne effect. An excellent example in psychotherapy research is

Luborsky's (1996) "symptom-context" technique for studying psychological symptoms as they arise in psychotherapy. The technique involves defining a temporal and interpersonal context in which to examine symptoms, identifying a comparison context in which the symptoms do not appear, analyzing at the individual level of analysis, and then aggregating the results. Using this technique, Luborsky demonstrated that symptom emergence almost always occurs in the context of the activation of what he calls an individual's "core conflictual relationship theme." While it took Luborsky many years to collect data on just seven individuals, the *PCSP* format would permit much more rapid collection of data on symptom contexts.

Study Micro-Episodes of Behavior and Process-Outcome Relationships

The *PCSP* journal is ideal for analyzing microepisodes of behavior and process–outcome relationships. The inclusion of a case formulation in each published case study means that a set of testable hypotheses about the client being treated is present. Those hypotheses could be tested in numerous ways. For example, interventions that are consistent versus inconsistent with case formulations have been shown to have effects on patient responses (Silberschatz & Curtis, 1993) and therapeutic outcomes (Silberschatz, Fretter, & Curtis, 1986). Replication of these studies is needed and could be facilitated by *PCSP* if transcripts were included as attachments to the case studies. The inclusion of plausible alternative hypotheses, as Curtis and Silberschatz (2007) recommend when constructing a case formulation using their Plan Formulation Method, would strengthen these studies. Edelson (1984; 1986) has also made a case for the scientific value of including alternate explanations in case study presentations, as later implemented by Fridhandler and Eells (1999).

Use PCSP to Develop Theory

A series of *PCSP* case studies could provide a format for developing theory about psychotherapy. Stiles (2003) describes a program of research involving a sequence of case studies that have helped develop and refine his assimilation theory of psychotherapy change. His view is that science is a process of comparing ideas (theories) with observations, and that theories improve as their fit with observations improves. He also makes clear that he views psychotherapy theories as practical tools to aid the therapist, not as context-independent universalities. A series of case studies provides an opportunity for conducting the comparisons to which he refers. For example, he and his colleagues documented in early case studies that therapeutic insights were accompanied by intense but mixed emotion. Consequently, they included an enhanced role for affective features in a revised theory of therapeutic change. Molenaar and Valsiner (2005) present an example of theory building using series of case studies involving time-series analyses.

Theory testing is often viewed as accomplished best through null hypothesis testing in the inter-individual framework. It may be too readily overlooked, however, that such testing never yields information about the probability the theory is correct in light of the data collected, but rather yields information about the probability that a null hypothesis is true (Krueger, 2001).

Analyzing series of case studies to build theory arguably provides more compelling and direct information about the theory itself and how well observations fit it.

Use PCSP to Study “Pure Cases”

The psychotherapy literature may benefit from greater study of “pure” cases in the Lewinian sense of the word. These are cases that best capture the phenomena under investigation regardless of whether the context is representative or not. Examples might be a psychotherapy analogue study that superbly captures the realism of actual psychotherapy while allowing for manipulation of experimental variables in ways that actual therapy does not. A recent study of psychotherapy case formulation expertise by Eells and colleagues (2005), which systematically varied case vignettes, represents such an intent. *PCSP* would be an excellent venue for presenting and critiquing “pure” case studies.

CONCLUSIONS

Single subject research can play a critical role in generating and generalizing knowledge about psychotherapy. Case studies are not simply stepping stones to more critical tests of theory (Flyvbjerg, 2006). I have demonstrated logical flaws involved in making inferences about individual psychological functioning when the method one uses permits only generalized inferences about groups, not the individuals comprising the groups. In contrasting Lewin’s conceptualization of the Aristotelian and Galileian modes of thinking in psychology, further logical flaws in casting single case research into a second class status were identified. It appears profitable to practice methodological pluralism when attempting to understand factors involved in psychotherapy process and outcome. Group and individual methods can mutually inform each other. The database being constructed by *PCSP* offers many opportunities to capitalize on the potential in single subject research.

In 1769, Giovanni Morgagni published *De Sedibus et Causis Morborum per Anatomen* (*The Seats and Causes of Disease Investigated by Anatomy*). Morgagni’s work is a compilation of over 700 well-indexed clinical case histories, each linking a patient’s symptom presentation to a report of pathology found at autopsy and any relevant experiments that had been conducted. *De Sedibus* was a remarkable achievement in the history of medicine in that it firmly established the anatomical concept of physical disease, which has developed further to include cellular and subcellular structures. An eighteenth century physician using *De Sedibus* to treat a patient could use the index to look up his patient’s symptoms, which could be cross-referenced to a list of pathological processes that may be involved. Morgagni also established the clinicopathological method of medical research, in which correspondences are examined between a patient’s symptoms and underlying pathology revealed at autopsy.

Is it possible to construct a psychological *De Sedibus*? If so, *PCSP* is a step in that direction. Although Morgagni had the advantage of postmortem anatomical results to correlate

with symptoms, his undertaking must have appeared no less daunting in its time than a corresponding undertaking to create a psychological *De Sedibus* appears now.

REFERENCES

- Bakan, D. (1954). A generalization of Sidman's results on group and individual functions, and a criterion. *Psychological Bulletin*, *51*, 63-64.
- Bakan, D. (1967). *On method: Toward a reconstruction of psychological investigation*. San Francisco: Jossey-Bass Inc., Publishers.
- Chase, W. G., & Simon, H. A. (1973). The mind's eye in chess. In W. G. Chase (Ed.), *Visual information processing* (pp. 215-281). New York: Academic Press.
- Clark, D. M., & Wells, A. (1995). A cognitive model of social phobia. In R. G. Heimberg & M. R. Liebowitz (Eds.), *Social phobia: Diagnosis, assessment, and treatment*. (pp. 69-93). New York, NY, US: Guilford Press.
- Cronbach, L. J. (1957). The two disciplines of scientific psychology. *American Psychologist*, *12*, 671-684.
- Curtis, J. T., & Silberschatz, G. (2007). Plan formulation method. In T. D. Eells (Ed.), *Handbook of psychotherapy case formulation* (2nd ed., pp. 198-220). New York: Guilford Press.
- Edelson, M. (1984). *Hypothesis and evidence in psychoanalysis*. Chicago: University of Chicago Press.
- Edelson, M. (1986). The evidential value of the psychoanalyst's clinical data. *Behavioral and Brain Sciences*, *9*, 232-234.
- Eells, T. D., Lombart, K. G., Kendjelic, E. M., Turner, L. C., & Lucas, C. (2005). The quality of psychotherapy case formulations: A comparison of expert, experienced, and novice cognitive-behavioral and psychodynamic therapists. *Journal of Consulting & Clinical Psychology*, *73*, 579-589.
- Flyvbjerg, B. (2006). Five misunderstandings about case-study research. *Qualitative Inquiry*, *12*, 219-245.
- Fridhandler, B., Eells, T. D., & Horowitz, M. (1999). Psychoanalytic explanation of pathological grief: Scientific observation of a single case. *Psychoanalytic Psychology*, *16*(1), 34-57.
- Gigerenzer, G. (1991). From tools to theories: A heuristic of discovery in cognitive psychology. *Psychological Review*, *98*, 254-267.
- Gould, S. J. (1996). *Full house: The spread of excellence from Plato to Darwin*. New York: Three Rivers Press.
- Kelley, H. H. (1967). Attribution theory in social psychology. *Nebraska symposium on motivation*, 192-240.
- Kelley, H. H. (1973). The processes of causal attribution. *American Psychologist*, *28*, 107-128.
- Koerner, K. (2007). Case formulation in dialectical behavior therapy for borderline personality disorder. In T. D. Eells (Ed.), *Handbook of Psychotherapy Case Formulation* (2nd ed.). New York: Guilford Press.
- Kraemer, H. C. (1978). Individual and ecological correlation in a general context. *Behavioral Science*, *23*, 67-72.
- Krueger, J. (2001). Null hypothesis significance testing. *American Psychologist*, *56*, 16-26.

- Kruger, L., Daston, L. J., & Heidelberger, M. (Eds.). (1987). *The probabilistic revolution, Vol. 1: Ideas in history; Vol. 2: Ideas in the sciences* (Vol. 1 & 2). Cambridge MA: MIT Press.
- Lambert, M. J., & Ogles, B. M. (2004). The efficacy and effectiveness of psychotherapy. In M. J. Lambert (Ed.), *Bergin and Garfield's Handbook of Psychotherapy and Behavior Change* (pp. 139-193). New York: John Wiley & Sons, Inc.
- Lewin, K. (1931). The conflict between Aristotelian and Galileian modes of thought in contemporary psychology. *Journal of General Psychology* 5, 141-177.
- Luborsky, L. (1996). *The symptom-context method: Symptoms as opportunities in psychotherapy*. Washington, DC: American Psychological Association.
- McArthur, L. A. (1972). The how and what of why: Some determinants and consequences of causal attribution. *Journal of Personality and Social Psychology*, 22(2), 171-193.
- Merton, R. (1989). *Serendipity: Accidental discoveries in science*. New York: Wiley.
- Molenaar, P. C. M., & Valsiner, J. (2005). How generalization works through the single case: A simple idiographic process analysis of an individual psychotherapy [Electronic Version]. *International Journal of Idiographic Science*. Retrieved September 3, 2006 from <http://www.valsiner.com/articles/molenvals.htm>.
- Morgan, D. L., & Morgan, R. K. (2001). Single-participant research design: Bringing science to managed care. *American Psychologist*, 56(2), 119-127.
- Morrison, D. E., & Henkel, R. E. (1970). *The significance test controversy*. Chicago: Aldine.
- Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.
- Nezu, A. M., Nezu, C. M., & Lombardo, E. R. (2004). *Cognitive-behavioral case formulation and treatment design: A problem-solving approach*. New York: Springer Publishing Co.
- Okiishi, J., Lambert, M. J., Nielsen, S. L., & Ogles, B. M. (2003). Waiting for supershrink: An empirical analysis of therapist effects. *Clinical Psychology & Psychotherapy*, 10, 361-373.
- Persons, J. B., & Tompkins, M. A. (2007). Cognitive-behavioral case formulation. In T. D. Eells (Ed.), *Handbook of Psychotherapy Case Formulation* (2nd ed.). New York: Guilford Press.
- Popper, K. R. (1959). *The logic of scientific discovery*. Oxford, England: Basic Books.
- Ricks, D. F. (1974). Supershrink: Methods of a therapist judged successful on the basis of adult outcomes of adolescent patients. In D. F. Ricks, M. Roff & A. Thomas (Eds.), *Life history research in psychopathology* (pp. 275-297). Minneapolis: University of Minnesota Press.
- Roethlisberger, F. J., & Dickson, W. J. (1939). *Management and the worker*. Cambridge, MA: Harvard University Press.
- Schacter, D. (2001). *The seven sins of memory: How the mind forgets and remembers*: Houghton, Mifflin.
- Sensky, T., Turkington, D., Kingdon, D., Scott, J. L., Scott, J., Siddler, R., et al. (2000). A randomized controlled trial of cognitive-behavioral therapy for persistent symptoms in schizophrenia resistant to medication. *Archives of General Psychiatry*, 57, 165-172.
- Sidman, M. (1952). A note on functional relations obtained from group data. *Psychological Bulletin*, 49, 263-269.

- Silberschatz, G., & Curtis, J. T. (1993). Measuring the therapist's impact on the patient's therapeutic progress. *Journal of Consulting and Clinical Psychology, 61*, 403-411.
- Silberschatz, G., Fretter, P. B., & Curtis, J. T. (1986). How do interpretations influence the process of psychotherapy? *Journal of Consulting and Clinical Psychology, 54*, 646-652.
- Skinner, B. F. (1956). A case history in scientific method. *American Psychologist, 221-233*.
- Stiles, W. B. (2003). When is a case study scientific research? *Psychotherapy Bulletin, 38*, 6-11.
- Stiles, W. B., Leach, C., Barkham, M., Lucock, M., Iveson, S., Shapiro, D. A., et al. (2003). Early sudden gains in psychotherapy under routine clinic conditions: Practice-based evidence. *Journal of Consulting & Clinical Psychology, 71*(1), 14-21.
- Tang, T. Z., & DeRubeis, R. J. (1999). Sudden gains and critical sessions in cognitive-behavioral therapy for depression. *Journal of Consulting and Clinical Psychology, 67*(6), 894-904.
- Thorngate, W. (1986). The production, detection, and explanation of behavioral patterns. In J. Valsiner (Ed.), *The individual subject and scientific psychology* (pp. 71-93). New York: Plenum Press.
- Tukey, D. D., & Borgida, E. (1983). An intrasubject approach to causal attribution. *Journal of Personality, 51*, 137-151.
- Valsiner, J. (1986a). Between groups and individuals: Psychologists' and laypersons' interpretations of correlational findings. In J. Valsiner (Ed.), *The individual subject and scientific psychology* (pp. 113-151). New York: Plenum Press.
- Valsiner, J. (1986b). Different perspectives on individual-based generalizations in psychology. In J. Valsiner (Ed.), *The individual subject and scientific psychology* (pp. 391-404). New York: Plenum Press.
- Wittgenstein, L. (1953). *Philosophical investigations* (G. E. M. Anscombe, Trans.). New York: Macmillan Publishing Company.
- Wolfe, B. E. (2005). *Understanding and treating anxiety disorders: An integrative approach to healing the wounded self*: American Psychological Association.