

# 2

## WHAT QUALIFIES AS RESEARCH ON WHICH TO JUDGE EFFECTIVE PRACTICE?

### Case Studies

*William B. Stiles*

In this position paper, I first discuss how scientific research provides quality control on theory. Then I argue that, for research on psychotherapy, case studies offer an alternative strategy that is as valuable as statistical hypothesis testing. Clinical practice is based on theory—if not a formal, stated theory, then an implicit one drawn from lore, convention, and personal experience. Case studies have some distinct scientific advantages for quality control on the complex, nuanced, context-responsive aspects of

---

I thank Giancarlo Dimaggio, Hani Henry, Mikael Leiman, James K. Mosher, Katherine Osatuke, and Lisa M. Salvi for comments on drafts of this chapter.

psychotherapy and psychotherapy theories, and they may be more satisfying to clinicians.

### THEORY IS THE MAIN PRODUCT OF SCIENCE

Theories are ideas stated in words (or numbers, diagrams, or other signs). Any account or explanation of something could be considered a theory. Commonsense and folk accounts of psychological disturbances and of how people overcome emotional and interpersonal difficulties can be considered theories, though they may be simplistic, internally inconsistent, imprecise, or unrealistic ones. The expectations that people use in daily dealings with each other may be considered as implicit theories. However, to be examined scientifically, a theory must be stated explicitly. In addition, a good theory should be internally consistent, precise, general, and realistic (cf., Levins, 1968).

A statement can be considered as accurate or true if it matches our observations of things or events, that is, if it matches what we see, hear, and feel. Of course, words and things are not the same stuff, but experience can be a common denominator, insofar as both statements and events are experienced. That is, I suggest the experience of an accurate or true statement corresponds in some way to the experience of observing the event it describes. This may be called an experiential correspondence theory of truth (Stiles, 1981, 2003). Such statements may be considered as facts if, additionally, there is agreement—social consensus—that they are accurate. In science, as in law, if people do not agree, then fact has not been established. A good theory, then, is one consistent with the facts, that is, with agreed descriptions of observations.

Research provides quality control on ideas by systematically producing observations and comparing them with the theories. The observations change the ideas and the theories. They may confirm or disconfirm a theory, or, more modestly, strengthen or weaken it. More often, however, the change involves extending, elaborating, refining, modifying, or qualifying the theory.

New observations may be said to permeate the theory. This is a diffusion metaphor; particles of observation spread through theoretical interstices. The ideas change to fit the observations, and aspects of the observations become part of the theory. The theory may be explained differently, for example, using different words that accommodate the new observations along with the previous ones or using the new observations as illustrations. Thus, the theory is modified by the observations to become more general, more precise, and more realistic. Darwin's (1859) theory of the origin of species by natural selection continues to grow as new observations elaborate and refine it. It was extended, for example, by the suggestion that the different nutritional values of the similarly sized bright red berries of dogwood trees (high in lipids and

other nutrients) and holly trees (mostly worthless fiber) in eastern deciduous forest may reflect coevolution with different populations of birds (respectively, autumn migrants who take dogwood berries on their way south and local residents who take holly berries in late winter after other resources are exhausted; Stiles, 1980).

Through research, then, observations accumulate in theories. New research results permeate the theory, but earlier thinking and results are retained. The diffusion metaphor offers an alternative to the brick wall metaphor for how science is cumulative. That is, understanding grows not by building a theoretical edifice, stacking fact upon fact, but rather by infusing observations that elaborate and change a theory in subtle (and sometimes not so subtle) ways. A living theory must be able to change, to accommodate this continual infusion of new observations; an impermeable theory is scientifically dead. Thus, permeability is a virtue in theories and in scientists (Stiles, 1993, 2003).

Permeability is a generalization of the traditionally acknowledged virtue of falsifiability. If theories or theoretical tenets were dichotomously true or false, permeability would be equivalent to falsifiability, insofar as the only change that evidence could make would be falsification (cf., Popper, 1934/1959). The concept of permeability, more realistically, I think, suggests that a theory is an approximation that can gradually change to represent scientists' experience more accurately as it accumulates observations.

If theories are informal or implicit, they may remain impermeable to evidence, insofar as the inconsistencies with observations are not exposed. Alternatively, informal theories may be too permeable, changing in response to each new observation and thus undermining any accumulation of understanding. Good theories must balance permeability with coherence and comprehensiveness, respecting previous observations while attending to new ones.

## STATISTICAL HYPOTHESIS TESTING VERSUS CASE STUDIES

I meant the foregoing to characterize both hypothesis testing and case studies. Both are strategies of empirical, scientific research, and both provide quality control on theory. Both yield observations that permeate theory, and both inform evidence-based practice (EBP), but they are different strategies.

The statistical hypothesis-testing strategy is to derive one or a few statements from a theory and compare each statement with many observations. If the observations tend to match the statement (in the investigator's experience, as conveyed to readers of the report), then people's confidence in the statement is substantially increased (e.g., not due to chance,  $p < .05$ ). This yields a small increment of confidence in the theory as a whole. For example, in the National Institute of Mental Health (NIMH) Treatment of Depression Collaborative Research Program (Elkin et al., 1989), clients'

scores on the Hamilton Rating Scale for Depression (HRSD) showed a statistically significant decrease across 16 sessions of manual-driven interpersonal therapy (IPT). This observation substantially increased confidence in the statement that clients' HRSD scores tend to decrease across 16-session IPT. The finding also added a small increment of confidence to the version of IPT theory from which it was derived (e.g., Klerman, Weissman, Rounsaville, & Chevron, 1984).

The case study strategy is to compare many theoretically based statements with correspondingly many observations. It does this by describing the case observations in theoretical terms. At issue is how well the theory describes details of the case (experiential correspondence of theory and observation). For reasons familiar to people trained in psychological research (selective sampling, low power, investigator biases, etc.), the consequent change in confidence in any one statement may be small, but because many statements are examined, the gain in confidence in the theory may be as large as from a statistical hypothesis-testing study. Campbell (1979) described this as analogous to the multiple degrees of freedom in a statistical hypothesis-testing study.

For example, the assimilation model (Stiles, 2002; Stiles et al., 1990) offers an account of how, in successful psychotherapy, clients assimilate problematic experiences through a sequence of stages. The problem moves from being warded off or dissociated through emerging and becoming understood to being worked through and mastered. The case of Fatima, a refugee (Varvin & Stiles, 1999), helped elaborate the warded-off stage, illustrating how warded-off material could appear initially as film-like memories—in Fatima's case, traumatic memories surrounding the birth and death of a daughter while Fatima was a political prisoner. This case study described how the memories emerged and were assimilated, at least partially. The film-like replaying of warded-off memories was consistent with the theory but not explicitly part of it previously. Incorporating these observations helped point toward similar manifestations of warded-off material in other cases of trauma. The consistency added a small increment of confidence in the theory, whereas the new observations extended it.

Because case studies do not focus on particular variables or hypotheses, the results of case studies are not in the form of decontextualized conclusions. The improvement in generality, precision, or realism is typically spread across the theory rather than concentrated in a sentence. The logic of case studies thus differs from the logic of  $n = 1$  designs, in which one or a few targeted dependent variables are examined over time as a function of the introduction or removal of independent variables. Unlike case studies,  $n = 1$  designs are meant to yield specific conclusions, stated in terms of the targeted variables.

A few systematically analyzed cases that match a theory in precise or unexpected detail may give people considerable confidence in the theory as a whole, even though each component assertion may remain tentative and uncertain when considered separately. Classic examples of such cases include

Dora for psychoanalysis (Freud, 1905/1953), Little Albert for behaviorism (Watson & Rayner, 1920), and Dibs for nondirective play therapy (Axline, 1964). I think that the degrees-of-freedom logic helps explain why such studies have had such impact. That is, readers were impressed because the studies reported many relevant observations on the cases (details, sequences, context), including some that were consistent with the theory but contrary to intuition or popular wisdom. It is worth noting that, although these cases generally tended to fit a theory that had been articulated previously, they also added or modified details, extending and enriching the theory, rather than merely illustrating it.

The statistical hypothesis-testing strategy can be problematic for studying psychotherapy. For statistical power, hypothesis-testing research must study common features. Common or recurring features of clinical cases are often artificial (e.g., Likert scale responses), global (e.g., efficacy of a treatment), or trivial. Single statements (hypotheses) out of context do not do justice to clinical theory as applied in practice, where it must accommodate variations in people, settings, and circumstances. Clinicians know that an in-session process is so full of nuances and responsive adjustments that researchers' labels or simple descriptions of global variables (e.g., "interpersonal therapy") do not adequately represent clinical reality. Consequently, such research results (increased confidence in isolated statements), even when they are positive, often fail to interest clinicians.

#### CASE STUDIES TRIANGULATE RATHER THAN REPLICATE

Case studies deal with the perennial tension between generality and uniqueness differently than hypothesis testing as each case includes details not shared with other cases. Statistical hypothesis testing seeks replication, and distinct features are often regarded as irrelevant or as errors in a statistical sense. In contrast, case studies use triangulation, considering distinct and novel features as explicitly informative (see Rosenwald's theory of multiple case research, 1988). Triangulation is a surveying metaphor that refers to the geometrical possibility of fixing a point in space by viewing it from two other locations. Distinctive features inform our understanding of the broader phenomenon of interest. For example, Fatima's film-like memories of her daughter's birth and death (Varvin & Stiles, 1999) offered a new perspective on warded-off material that complemented the assimilation model's previous accounts. Of course, exact replications are never possible, and even traditional replication studies involve alterations or extensions, so what is replicated is the interpretation rather than the observation.

In the Indian parable, as retold by the American poet John Godfrey Saxe (1816–1887), six blind men were each led to a different part of an elephant and asked to describe the beast. The man who felt its side said the

elephant was like a wall; the man who felt its tusk, like a spear; the man who felt its trunk, a snake; the man who felt its leg, a tree; the man who felt its ear, a fan; and the man who felt its tail, a rope. Although the men in the story refused to agree or listen to each other, the point of the parable, like multiple case research, is that the beast has many aspects, and understanding it demands multiple, diverse perspectives. A logic that restricts attention to common features is likely to narrow an account of psychotherapy into oblivion or banality. Different cases and different perspectives may be expected to yield different interpretations, and it is the task of theorists to reconcile and integrate them. In this way, allowing the unique aspects of each case to permeate the theory increases the theory's generality.

### APPLY THE CASE TO THE THEORY, NOT ONLY THE THEORY TO THE CASE

The logic of permeability suggests that a scientific case study is meant to change the theory, not to understand the case. As in any scientific research, observation is the authority, and a theory that does not converge with observation must be changed. Thus, an investigator must have the confidence to modify the theory—to extend its scope, to change its expression, to add details, and so forth.

A scientific case study thus contrasts with the clinical use of theory, in which the point is to apply theories to understand clinical phenomena. Many case studies miss this distinction, adhering to previously stated theory and ignoring, discounting, or distorting observations that fail to fit. Using the theory to understand the case is an essential first step in research, showing the ways in which the current theoretical account matches the new observations. But investigators must be willing to focus on features of the case that go beyond or differ from the current theoretical account, to turn the observations back on the theory to improve it. A respectful attitude toward theory has a place in clinical applications, insofar as clinicians, and the rest of us, often fail to appreciate phenomena to which the theories are pointing, so that crediting the theory above our own initial impressions or prejudices may open us to things we had overlooked. But case studies that merely apply theories, without putting them at risk of alteration, do not make a scientific contribution.

The logic of permeability also demands a technology of reconciliation. Theories should be internally consistent—logically coherent—as well as consistent with observations. Modifications to theory from different case studies (or any sort of studies) must be reconciled. Changes made to help understand a new case are unhelpful if they discount or distort observations on previous cases. Thus, systematic case studies demand equally systematic conceptual reviews in which inferences on the basis of different cases are compared and reconciled. I suspect that such reconciliation is best accomplished in a dia-

logical process of statement and response, as well as logical analysis. Perhaps new techniques are needed here, such as systematic ways to reconcile versions of a theory that derive from encompassing different observations.

### ADVANTAGES AND DISADVANTAGES

Practitioners have been chronically unhappy with psychotherapy research (e.g., Morrow-Bradley & Elliott, 1986; Talley, Strupp, & Butler, 1994). They often find statistical hypothesis studies narrow, tedious, and too decontextualized to be assimilated into their practice. Many difficulties have contributed to the research-practice gap, of course, but case studies might address some of them.

Case studies can address the complexity and subtlety of psychotherapy. Triangulation allows a case study to encompass nuances and unique context, making research seem more realistic to practitioners while also making productive use of case-to-case and session-to-session variations. By incorporating clients' and therapists' individuality, case studies may also be relatively easily integrated with practitioners' humanitarian values. Case studies may be better suited than hypothesis-testing research to the complex theories practitioners use and the highly contextual material practitioners encounter. Ethical and practical constraints often prevent scientists from exercising the sort of control needed for hypothesis-testing research on psychotherapy, whereas case study research can use practice-based clinical presentations and clinical intervention. In growing recognition of such advantages, several relatively new journals offer outlets for case studies, including *Pragmatic Case Studies in Psychotherapy*, *Journal of Clinical Psychology: In Session*, and *Clinical Case Studies*.

In addition to speaking to practitioners as consumers of research, case studies may offer new opportunities for practitioners to conduct useful research. Practitioners who lack the resources to conduct clinical trials may be able to conduct case studies. Anyone who has practiced realizes that, in addition to its potential for healing, psychotherapy can be a marvelous laboratory for observing human experience and behavior. Psychotherapists regularly see aspects of people that others seldom or never see. They thus have an exceptional opportunity to make empirical observations that could bear directly on psychotherapy theory. In principle, case studies offer a way to use this opportunity to improve theory.

Case studies also have disadvantages. Campbell's (1979) degrees-of-freedom argument (i.e., that a study of a single case involves many observations) addresses the argument that case studies lack power, but it does not overcome other familiar objections. Such objections include selective sampling (choice of cases), reliability questions (trustworthiness of observations), the imprecision of measurement, the lack of standard vocabulary, and investigator biases. Case studies also may be more difficult to write about, as

they may require more detailed descriptions and rely less on the standardized language of research reports.

The lack of succinct conclusions can make case studies profoundly puzzling to readers used to the results of hypothesis-testing research that can be stated in a sentence. The lack of statistical independence and the associated issues (e.g., collinearity and compounded alpha levels) are not problematic, as independence is not assumed. Indeed, interrelations among observations may make relevant and interesting contributions to descriptions. Different phrases may be needed to characterize how case studies support or fail to support a theory, and questions of validity and alternative explanations must be cast differently when the value of the study depends on many tentative empirical statements, rather than one or a few firm ones. Case studies make the theory-building logic of science more explicit. By offering no context-free conclusions, the case study logic makes it harder to maintain an illusion of context-free knowledge. That is, case studies demand explicit theories, and in my view this is a virtue.

### CONCLUSIONS

Both case studies and hypothesis-testing research can provide quality control on theories, which is the main point of scientific research. Both strategies yield observations that permeate the theories, making them more general, precise, and realistic. Case studies use a strategy that is different from hypothesis testing, addressing many theoretical issues in the same study, rather than focusing on only one or a few. Case studies incorporate unique features of cases, emphasizing triangulation rather than replication. Despite familiar drawbacks, case studies have distinctive advantages for research on psychotherapy, particularly their ability to study multifaceted phenomena in context. In summary, case studies are well qualified to provide evidence on the theories that underlie the effective practice of psychotherapy.

## References

- Axline, V. (1964). *Dibs: In search of self*. Boston: Houghton-Mifflin.
- Campbell, D. T. (1979). "Degrees of freedom" and the case study. In T. D. Cook & C. S. Reichardt (Eds.), *Qualitative and quantitative methods in evaluation research* (pp. 49-67). Beverly Hills, CA: Sage.
- Darwin, C. (1859). *On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life*. London: J. Murray.
- Elkin, I., Shea, M. T., Watkins, J. T., Imber, S. D., Sotsky, S. M., Collins, J. F., Glass, D. R., Pilkonis, P. A., Leber, W. R., Docherty, J. P., Fiester, S. J., & Parloff, M. B. (1989). National Institute of Mental Health Treatment of Depression Collaborative Research Program: General effectiveness of treatments. *Archives of General Psychiatry*, *46*, 971-982.
- Freud, S. (1953). Fragment of an analysis of a case of hysteria. In J. Strachey (Ed. & Trans.), *The standard edition of the complete psychological works of Sigmund Freud, Vol. 7* (pp. 3-122). London: Hogarth. (Original work published 1905).
- Klerman, G. L., Weissman, M. M., Rounsaville, B. J., & Chevron, E. S. (1984). *Interpersonal psychotherapy of depression*. New York: Basic Books.
- Levins, R. (1968). *Evolution in changing environments: Some theoretical explorations*. Princeton: Princeton University Press.
- Morrow-Bradley, C., & Elliott, R. (1986). The utilization of psychotherapy research by practicing psychotherapists. *American Psychologist*, *41*, 188-197.
- Popper, K. (1959). *The logic of scientific discovery*. New York: Basic Books. (Original work published 1934).
- Rosenwald, G. C. (1988). A theory of multiple case research. *Journal of Personality*, *56*, 239-264.
- Stiles, E. W. (1980). Patterns of fruit presentation and seed dispersal in bird-disseminated woody plants in the eastern deciduous forest. *The American Naturalist*, *116*, 670-688.
- Stiles, W. B. (1981). Science, experience, and truth: A conversation with myself. *Teaching of Psychology*, *8*, 227-230.
- Stiles, W. B., Elliott, R., Llewelyn, S. P., Firth-Cozens, J. A., Margison, F. R., Shapiro, D. A., & Hardy, G. (1990). Assimilation of problematic experiences by clients in psychotherapy. *Psychotherapy*, *27*, 411-420.
- Stiles, W. B. (1993). Quality control in qualitative research. *Clinical Psychology Review*, *13*, 593-618.
- Stiles, W. B. (2002). Assimilation of problematic experiences. In J. C. Norcross (Ed.), *Psychotherapy relationships that work: Therapist contributions and responsiveness to patients* (pp. 357-365). New York: Oxford University Press.
- Stiles, W. B. (2003). Qualitative research: Evaluating the process and the product. In S. P. Llewelyn & P. Kennedy (Eds.), *Handbook of clinical health psychology* (pp. 477-499). London: Wiley.
- Talley, P. F., Strupp, H. H., & Butler, S. F. (Eds.) (1994). *Psychotherapy research and practice: Bridging the gap*. New York: Basic Books.
- Varvin, S., & Stiles, W. B. (1999). Emergence of severe traumatic experiences: An assimilation analysis of psychoanalytic therapy with a political refugee. *Psychotherapy Research*, *9*, 381-404.
- Watson, J. B., & Rayner, R. (1920). Conditioned emotional reactions. *Journal of Experimental Psychology*, *3*, 1-14.