Commentary on *Logical Operations in Theory-Building Case Studies* by William Stiles

What Can We Learn from Case Studies?
More than Most Psychologists Realize

MICHAEL A. WESTERMAN [a,b]

[a] Department of Psychology, New York University

[b] Correspondence should be sent to Michael A. Westerman, Department of Psychology, New York University, 715 Broadway, 2nd Floor, New York, New York, 10003.
Email: michael.westerman@nyu.edu.

ABSTRACT

Stiles (2009) was on target when he argued that theory-building case studies have much more to offer than most psychologists realize. In this analysis of Stiles' position, I focus on comparisons he drew between theory-building case studies and hypothesis-testing investigations with regard to (a) sensitivity to context, and (b) evaluating theories. Stiles' suggestion that case studies are context-sensitive whereas hypothesis-testing studies are not misses some of the actual complexity of that issue because some hypothesis-testing studies reflect genuine appreciation of context. I identify specific differences that exist between the two research approaches with respect to sensitivity to context, including particular ways in which case studies are better with respect to this concern. Stiles' comments about how case studies can contribute to evaluating theories are novel, convincing, and very important. The key point Stiles developed is that case-study research is especially well suited for theory evaluation because it allows investigators to compare observations to the multiple tenets of a theory at the same time, whereas hypothesis-testing studies test single tenets of theories against observations. I illustrate Stiles' model with an example from a set of case studies of my own on the theory of interpersonal defense.

Keywords: case studies; hypothesis-testing investigations; evaluating theories; context sensitivity

---

Stiles (2009) offered a fascinating examination of the epistemology of case-study research. He maintained that a particular approach to such research, which he calls "theory-building case studies," represents an important method for advancing understanding of psychotherapy. In my opinion, Stiles' account is on the mark. Case-study research of the kind Stiles advocates should be accorded the status of scientific research, even though up until now many psychologists probably would disagree with this claim.

Although I agree with the main thrust of Stiles' position, I also believe that a close examination—including critical consideration—of his ideas is in order. We need to follow Stiles' lead in revisiting basic questions about epistemology and the philosophy of science as they relate to case studies. I will offer a number of thoughts along these lines, taking as my point of departure the comparisons Stiles drew between hypothesis-testing research and theory-building case studies. My comments will present some differences between my understanding of these

issues and the position Stiles put forward. They will also highlight, elaborate upon, and offer an illustration of some of the points Stiles has advanced that I agree with and believe to be especially important.

COMPARING THEORY-BUILDING CASE STUDIES AND HYPOTHESIS-TESTING RESEARCH

Probably the most important point Stiles makes in his comparison between hypothesis-testing studies and case-study research is that the two are similar in the sense that both can be used for evaluating theories, so long as case studies are pursued as theory-building investigations. In other words, as Stiles (2009, p. 12) indicated, case studies can play a role in justifying theories; they are not limited to the discovery phase of research, though many psychologists have maintained that they only provide a basis for generating theories, not evaluating them. Note, also, that by focusing on the role case studies can play with respect to the assessment and development of theories, Stiles clearly staked out a position according to which case-study research can provide knowledge of general applicability, not only understanding of the particular cases examined. Therefore, his advocacy of case studies differs from the position taken by psychologists who believe that research on individual cases is important precisely because our understanding of psychological phenomena ultimately is case specific (e.g., Fishman, 1999).[1]

Stiles also argued that hypothesis-testing investigations and case studies are different in several respects. The key idea here is that a hypothesis-testing study tests a single tenet of a theory against observations, whereas a case study offers the opportunity for evaluating the theory itself, including its multiple tenets, against observations. According to Stiles, the two research approaches also differ because hypothesis-testing research focuses on commonalities across cases, whereas case studies can explore the unique features of individual cases as well as commonalities. In addition, Stiles maintained that the two approaches are different with respect to context sensitivity: cases studies can be sensitive to context, whereas hypothesis-testing studies are limited in this respect.

Context-Sensitive Research

I will begin my examination of these comparisons by considering the last point, sensitivity to context. In the article under consideration, Stiles (2009) made strong claims about how case studies are superior to hypothesis-testing investigations in this regard. For example, he argued that:

> ... case studies can bring contextually specific features into theory. By observing and describing the details of what particular individuals say and do and experience – and when they say, do, or experience it – case studies can address theories in ways denied to

---

[1] Fishman (1999) argued for developing case specific accounts based on his reading of the implications of pragmatism and hermeneutics. At a later point, I will offer some comments about how case studies can contribute to evaluating theories of general applicability (in support of Stiles' position) based on my own view of the implications of pragmatism and hermeneutics.

> hypothesis testing, which is restricted to phenomena that can be quantified and counted and observed in frequencies large enough for inferential comparisons. Through case studies, then, the richness of empathy with another human being can be registered, and inhuman or mechanical theories can be challenged (p. 18).

In an earlier article, Stiles (2003 [reprinted as Appendix A in Stiles, 2009]) made some additional claims along these same lines:

> In contrast to statistical hypothesis-testing research, case studies characteristically yield results mainly in words rather than numbers, use empathy and personal understanding rather than detached observation, place observations in context rather than in isolation, focus on good examples rather than representative samples... (p. 7)

Stiles (2009) maintained that these differences are extremely important because psychotherapy phenomena are such that context-sensitive methods are crucial for advancing knowledge in this domain. I agree with this viewpoint. Indeed, I believe that context-sensitive research methods are crucial for advancing knowledge in most, if not all, areas of the discipline. Nevertheless, I disagree with the comparison drawn by Stiles between case studies and hypothesis-testing research when it comes to sensitivity to context. Specifically, I believe that hypothesis-testing studies have more to offer on this score than Stiles acknowledged, at least in the article of present interest or the other publication from which I just quoted (Stiles, 2003).

It may be that the roots of this difference of opinion can be traced to differences on basic issues regarding the philosophy of science. For present purposes, it is useful to offer some brief considerations at that level of analysis. Stiles and I both believe that research on psychotherapy should be an interpretive process of inquiry that is based on appreciation of context and meaning, but we hold this view for very different reasons. Stiles' position is based on his experiential correspondence theory of truth. According to that view, "a theory is a good one if people's experiences of the theoretical descriptions (i.e., the meanings of the descriptions to them) correspond with their experiences of observing the objects and events in the world" (Stiles, 2009, p. 10).

As I see it, this experiential correspondence theory is dualistic, like most of our philosophical tradition. The theory places signs and meanings on one side (private) and observable events on the other (public). To be precise, it might be more accurate to describe the theory as a monistic account that reflects one common variant of our dualistic tradition. According to this second reading, the theory focuses exclusively on the side of the private. The matching process that Stiles referred to checks correspondence between "experiences of the theoretical descriptions" and "experiences of observing the objects and events in the world." Note, however, that this monistic reading actually rests on dualist philosophy, because the latter is the source of the notion of a separable private domain.

By contrast, my support of interpretive methods of inquiry is based on a philosophical perspective which maintains that the private is always embedded in the world, not separable from it. This perspective is derived from the work of Merleau-Ponty (1962), Wittgenstein (1958), Heidegger (1962), and the American pragmatists (e.g., Dewey, 1896). The starting point idea for this perspective is the notion that the person is always already actively involved in the world of

meaningful practical activities. This idea, which Merleau-Ponty (1962) called "involved subjectivity," departs from the notion of a subject separate from the world, which is the cornerstone of our philosophical tradition. Involved subjectivity also differs from both of the monistic positions associated with the notion of a subject separate from the world – positions which maintain that the subject's experience *is* the world and positions which put forward the idea of a mechanical world that encompasses everything and leaves no room for the subject.

As compared to Stiles' experiential correspondence theory of truth, taking involved subjectivity as the starting point leads to a different understanding of why interpretation enters into the research process in psychology – even though both Stiles and I believe that interpretation should play a central role in psychological inquiry. As I see it, interpretation is central not because the meanings of human behavior refer to the subject's private experience or understanding of events, but precisely because those meanings always go beyond the individual subject to refer to his or her involvement in practices. Our "take on things" always refers beyond us to what we cannot say explicitly and what we do not fully know because it is embedded in our concrete participation in doing things.

This view of why our understanding is irreducibly interpretive, in turn, leads to a very different view of what the interpretive process is like and what is involved in taking context into account. Along with Stiles, I believe that psychological phenomena are much more than a concatenation of brute, meaningless events. We cannot dispense with, or explain away, meanings in our research efforts. In particular, the meanings of human behavior cannot be built up from meaning-free isolable behaviors, because the significance of individual behaviors always refers to the context of the practical activities of which they are a part.

But how do we understand the way in which meaning enters into the behavior-context relationship? In my opinion, individual behaviors have significance because of the role they play in the contexts in which they appear, but the behavior-context relationship does not refer to abstract, rarified, and private meanings (templates, rules, or schema) that lie "behind" concrete events. Rather, the behavior-context relationship refers to meanings that are in-the-world. Meanings irreducibly refer to the concrete shapes of the activities in which we are engaged. Therefore, psychological phenomena are meaningful in a way that it is fully indexical, that is, embedded in actual concrete situations. In order to understand psychological phenomena, we must place our observations in the meaningful contexts of what the people we are studying are doing. These contexts are more than "just" a set of specific behaviors, but they are not something separable (abstract ideas, templates, and so forth) from concrete activities.

It follows from these considerations that hypothesis-testing studies are misguided if they consider observations in isolation. In good measure, Stiles was correct when he criticized hypothesis-testing research for being insufficiently sensitive to context because much of that work suffers from considering observations in isolation. In fact, Stiles'(1988) work on therapist responsiveness identifies one of the key ways in which this problem appears in the area of psychotherapy research. In his work on responsiveness, Stiles pointed out how process-outcome studies typically focus on correlations based on frequency counts of particular therapist behaviors as if we can abstract these individual behaviors from the contexts in which they occur. This approach neglects how specific therapist bids relate to what is happening at the moment in

which they take place. Here I am in agreement with Stiles, but I disagree with the suggestion, which I believe comes across clearly in the article under consideration, that failing to be sensitive to context is a necessary characteristic of hypothesis-testing investigations.

Elsewhere, I have argued that the philosophical perspective of involved subjectivity leads to recognizing that systematic quantitative research, including hypothesis-testing studies, can provide a way to pursue inquiry that is interpretive and context sensitive (Westerman, 2004, 2006a, 2006b, in press). In fact, I have argued that it follows from the basic philosophical perspective that such research can make a unique contribution to efforts to learn about psychological phenomena, even though qualitative studies also have a great deal to offer. In brief, the main idea here is that because phenomena psychologists try to learn about are concretely meaningful, not meaningful in some abstract, rarified sense, it is important to concretely specify our theoretical ideas. The step in quantitative research that we typically refer to as "operationalizing" hypotheses provides an opportunity for doing just this, if we proceed along the lines of what I call "explicitly interpretive quantitative inquiry" and avoid the common pitfall of trying to operationalize ideas of interest in a manner aimed at explaining away meanings.

In quantitative research of the sort I advocate, coding schemes in observational studies concretely specify meaningful phenomena, not reduce them to tallies of brute behaviors, and experimental paradigms concretely specify hypothesized causal processes, because they are situations that help us understand concretely how those meaningful processes work. At one and the same time, explicitly interpretive quantitative research helps us learn about ideas concerning meaningful psychological phenomena and what those phenomena are like concretely.

To cite one example that illustrates these points: Relational codes in observational research on interpersonal interaction are often very useful. Rather than indexing individual behaviors viewed as such, these coding categories provide ways to characterize how one turn in an interaction relates to other turns. For example, Keenan and Schieffelin (1976) defined the coding categories, "collaboration" and "incorporation." In brief, a turn that is a collaboration addresses the same topic as preceding turns, but does not add new information to the exchange that goes beyond a possible affirmation or negation, whereas an incorporation addresses the same topic and adds something new to the interaction. Clearly, these behavior codes refer to meaningful processes (considerations about meanings are essential for determining what counts as a collaboration or an incorporation). At the same time, these codes concretely specify different ways in which one response can relate to another, and they do this based on careful, concrete descriptions (that make use of examples) of these phenomena.

It is interesting to note that in my publications about these ideas I have used certain terms that Stiles also used, but in a different way. As I see it, when employed in a helpful manner, coding schemes in observational research and paradigms in experimental investigations provide "good examples" of concretely meaningful phenomena, not representative samples of behavior. By contrast, as noted earlier, Stiles (2003, p. 15) suggested that case studies focus on good examples, whereas hypothesis-testing studies focus on representative samples.

If I am on the right track, then Stiles' article gives too little credit to hypothesis-testing research and it is important to consider a more positive viewpoint about what can be achieved by means of hypothesis-testing studies. Nevertheless, for three reasons, I do not want to put too fine a point on what might appear to be a fairly straightforward disagreement between my position and Stiles'. The first reason returns to something I said earlier. I believe that Stiles is on the mark when he takes hypothesis-testing research to task in the sense that his critique about failing to be context sensitive does apply to a great deal of research of that kind. Second, Stiles has made it clear in other publications that he recognizes that hypothesis-testing investigations can be conducted along the lines I have briefly pointed to here (Osatuke & Stiles, in press; Stiles, 2006). In fact, I believe that Stiles' own research on his assimilation model, including some hypothesis-testing studies Stiles has conducted, offers some of the best examples of explicitly interpretive quantitative research (see Osatuke & Stiles, in press).

The third reason is that, although I believe that hypothesis-testing research can be conducted in ways that are sensitive to context, I do not wish to suggest that when hypothesis-testing research is conducted in that manner there is no difference at all between how such research is sensitive to context and how case-study research can be context-sensitive. There are differences, even though it is a complicated matter to identify just what they are. I will attempt to do that in what follows.

Differences Related to Context Sensitivity

Although I do not believe that case studies and hypothesis-studies differ in the simple sense that the former approach is context-sensitive while the latter is not, there is an important respect in which the two approaches differ that is related to context sensitivity. For the most part, case-study research does not involve systematic measurement procedures, whereas this is an essential feature of hypothesis-testing studies.[2] Even when measurement is carried out as part of explicitly interpretive quantitative research, this does make for important differences between hypothesis-testing investigations and case studies.

For one thing, in case-study research, it is easier to focus on meaningful phenomena in ways that examine how those phenomena appear differently in different contexts. Explicitly interpretive hypothesis-testing studies are based on real appreciation of context, but that research strategy is not as flexible with regard to context as is the case-study approach. For example, imagine that an investigator conducting a case study is interested in the phenomenon of incorporation as he or she examines the transcript of a therapy session. Now imagine that a remark by the patient addresses the topic that is on the table and adds to what the therapist just said, but also imagine that this "addition" actually repeats what the patient said in several previous bouts of exchange during the same session. Finally, imagine that this phenomenon (adding apparently new material that actually represents mere repetitions) recurs multiple times during the session. In the process of working on a case study, there is a good chance that a

---

[2] This distinction does not always hold. Some case-study research does make use of formal measurement procedures. In fact, case studies on Stiles' assimilation model often make use of his APES scale, an ordinal scale that characterizes steps in the developmental process of integrating conflicting voices in psychotherapy (see Osatuke & Stiles, in press).

researcher would recognize this phenomenon. He or she might well realize that the apparent incorporations do not really add anything new at a more molar level of analysis – and that, therefore, they are not truly incorporations. By contrast, there is a real possibility that an investigator studying similar ideas by means of a hypothesis-testing investigation might fail to appreciate what occurred in the session, if his or her system for coding incorporations defined the phenomenon in terms of how a given turn related to the previous bid in the exchange.

This is certainly an important difference, but there are pros and cons on both sides of the ledger. The point I have just made identifies a real "plus" for case studies, but for some research purposes, there are considerable benefits to employing systematic measurement procedures because those procedures help us concretely specify phenomena of interest. To go back to my example about incorporations, although it may be true that there is a similarity between addressing the topic on the table but not adding anything new as compared to what was just said, on the one hand and, on the other hand, addressing the topic on the table and not adding anything new as compared to what was said over the course of an entire exchange, there certainly could be research situations in which we might want to know what concrete form this phenomenon typically takes. That is, we might want to know what would be a truly good concrete example.

To illustrate this point: Consider a hypothetical study on the development of discourse capabilities in young children. Although, conceivably, failures of incorporation could take either form just described, it might well be that, in fact, only the phenomenon defined in terms of how a given bid relates to the immediately preceding turn captures anything of significance when it comes to the interaction behavior of children of a certain young age – perhaps because interaction behaviors for young children simply do not cohere across spans of time. It could be that the issue of whether a child adds to what has been just said and also to what was said at earlier points in an exchange only becomes relevant at a later age. While we might recognize this shift in careful case studies of early child development, hypothesis-testing studies are better-suited for this research situation in which we would be trying to determine what the relevant concrete form is like. This point is very much in keeping with the thrust of Wittgenstein's (1958, § 215) question, "But isn't *the same* at least the same?" What appears to be "the same" from a logical or removed point of view may not be "the same" at all psychologically.

These considerations are related to another "plus" in favor of hypothesis-testing studies. I believe that Stiles was right when he argued that case studies can contribute to validating theories, but I think the strengths they have along those lines are different from the strengths hypothesis-testing studies have for justification-oriented investigations. In my opinion, hypothesis-testing research can be especially helpful for making systematic comparisons. By concretely specifying processes of interest, systematic measurement procedures make possible quantitative comparisons. Such comparisons can be very useful. The informal comparisons typically made as part of case-study research are subject to such problems as the tendency to give too much weight to striking occurrences of a phenomenon.

To be sure, systematic comparisons in hypothesis-testing research are not always helpful. If we employ the "wrong" coding system (that is, one that is not based on good examples) for assessing incorporation, we could come up with mistaken conclusions. Nevertheless, through an iterative process across multiple hypothesis-testing testing studies, we might very well arrive at a

truly helpful concrete specification of the phenomenon of interest and we may well succeed at learning valuable things about it by making careful systematic comparisons.

These points highlight the importance of considering the purpose of a given research project when evaluating different methods. In the article under consideration, Stiles made a valuable observation of this sort. I believe that Stiles was correct when he suggested that one of the strengths of case studies is that they are likely to lead to results with real clinical utility. In many situations, clinicians can make good use of results that focus on meaningful phenomena in a manner that recognizes that they can appear in different ways in different contexts (i.e., the kind of findings that case studies often provide), because clinicians can map such results to the array of diverse contexts that come up as they work with patients.

This is a "plus" for case studies, but it comes with qualifications. For example, while a case study investigator or a clinician might recognize that a patient is not really "incorporating" if he or she repeatedly adds to what was just said but in a way that does not really contribute anything new to the overall exchange, the investigator or clinician also might miss this. Quantitative studies could help therapists recognize the relevance of the more molar context when they think about phenomena related to incorporation.

To summarize what I have said about the issue of context sensitivity: No simple answer can be offered about whether case studies are better than hypothesis-testing investigations with regard to being sensitive to context. I think Stiles missed some of the complexities of this issue when he suggested in the paper under consideration that case studies are context-sensitive whereas hypothesis-testing studies are not, because the latter can reflect genuine appreciation of context when they are conducted as explicitly interpretive quantitative investigations. Nevertheless, I have also argued that the two research approaches are not the same when it comes to this issue. There are ways in which case studies are preferable with respect to appreciation of context – although it may not be possible to identify any particular "plusses" for case studies that always hold. I have suggested that any comparison we might try to make between the merits of case studies and hypothesis-testing research with respect to context sensitivity will depend on the purposes of the particular research project under consideration. Often, if the purpose relates to clinical utility, the balance will be in favor of case-study research.

EVALUATING THEORIES VERSUS EVALUATING TENETS OF THEORIES

Case studies have much to offer with regard to sensitivity to context, but comparisons between case-study research and hypothesis-testing investigations are complicated on this score. As I see it, the strongest case for the merits of case studies lies elsewhere. This brings us to what I believe is the most noteworthy contribution of Stiles' article (2009), the idea that hypothesis-testing studies test single tenets of theories against observations, whereas case studies provide a method for evaluating theories, including their multiple tenets, against observations. Stiles' emphasis on theories seems right on the mark to me, and I believe he is correct in arguing that case studies have a different and better relationship to theory-building than do hypothesis-testing studies.

An Example: The Theory Of Interpersonal Defense

I can illustrate how case studies and hypothesis-testing investigations differ with respect to testing complete theories by offering some brief comments about a program of research I have been pursuing for some time. I have been developing an interpersonal reconceptualization of defense processes called the theory of interpersonal defense (e.g., Westerman, 1998; Westerman & Steen, 2007, in press). Four of the key tenets of the theory are:

(a) <u>nature of defense processes:</u> processes of defense primarily involve patterns of interpersonal behavior,

(b) <u>discourse breaches/failures of coordination:</u> defensive interpersonal behavior patterns are characterized by recurring discourse breaches, or what I also refer to as failures of coordination,

(c) <u>negotiating conflicts:</u> these patterns represent attempts to negotiate conflicts between pursuing wished-for relationship outcomes and feared consequences that might occur if a person pursues his or her wished-for outcomes, and

(d) <u>"feed forward effects":</u> defensive patterns have what I call "feed forward effects," that is, they influence what subsequent events will occur in interactions.

Tenet (d) includes several more specific tenets. According to one of these, defensive behavior "works" in the sense that it reduces the likelihood that clear-cut instances of feared consequences will occur. However, according to other parts of the theory, it also has undesired feed forward effects, including making it less likely that clear-cut occurrences of wished-for outcomes will take place and more likely that negative outcomes will occur that are distinct from, but related to, feared consequences.[3] Finally, according to the theory, defensive behavior also leads to positive outcomes that are distinct from, but related to, wished-for outcomes.

Clearly, this is a complex theory with many parts to it. Moreover, the parts are interrelated (which would be true for the tenets of any theory). For example, the theory offers a characterization of the nature of defensive interpersonal behavior patterns (briefly summarized in tenets (b) and (c) above) that dovetails with the specific tenets about feed forward effects (tenet (d)). These two sets of claims must hang together, that is, the characterization of defensive behavior has to describe patterns that would lead to the kinds of effects that are described in other parts of the theory.

Until recently, I pursued two lines of inquiry related to the theory. Both lines involved hypothesis-testing studies. In one line, my colleagues and I conducted observational studies of psychotherapy process and examined patient in-session interpersonal behavior in terms of tenet (b) about discourse breaches/failures of coordination (Westerman & Foote, 1995; Westerman,

---

[3] Note that the theory differs from usual formulations of the self-fulfilling prophecy notion which claim that, in the long run, problematic interpersonal behavior actually leads to the very negative consequences a person was trying to avoid (see Dahmen & Westerman, 2007).

Foote, & Winston, 1995). In other research, I conducted experimental studies with colleagues on tenet (d), the feed forward claims of the theory (Dahmen & Westerman, 2007; Westerman & Prieto, 2006). Results from both lines of inquiry have supported the relevant tenets of the theory. The observational studies of therapy process have shown that patients' behavior can be reliably assessed in terms of coordination and that these assessments are related to outcome – and differentially related to outcome as a function of differences in therapeutic approaches – in ways that support the idea that coordination indexes defensive behavior. The experimental studies have provided preliminary support for all of the tenets about feed forward effects.

I recently began a new line of inquiry that involves conducting a small set of intensive case studies (Westerman, 2008). This line of inquiry has demonstrated to me that Stiles was right on target when he argued that case studies are especially well-suited for evaluating theories because they include comparing observations to the multiple tenets of a theory.

The case studies I have conducted on interpersonal defense theory (Westerman, 2008) illustrate Stiles' crucial point that this methodology allows for testing multiple hypotheses simultaneously, because all of the interpersonal phenomena predicted by interpersonal defense theory should be present in videotapes of actual therapy sessions. Specifically, I found that observations from each of the cases I studied simultaneously conformed to the several parts of the theory because: (1) defensive interpersonal behavior was characterized by recurring discourse breaches, (2) for each case, it was possible to understand how these recurring patterns represented attempts to pursue wished-for relationship outcomes while trying to avoid feared consequences that might result from pursuing the wishes, (3) defensive patterns actually led to the complex feed forward effects described by the theory (see tenet (d) above) in the patients' relationships with significant others besides the therapist, and (4) although these feed forward effects also occurred to some extent in the therapist's behavior in all of the cases, this was true to a lesser extent in some cases than in others, especially in the middle and late phases of treatment. Outcome was good in the cases with less pronounced feed forward effects, whereas it was poor in those cases in which the feed forward effects remained a prominent feature of therapist behavior throughout the course of treatment. The fact that observations from the case studies supported all of these predictions from interpersonal defense theory *at the same time* constitutes an important kind of support for the theory. It illustrates a unique way in which case studies can contribute to theory building.

To make this point in a more concrete fashion, I can offer a brief consideration of an excerpt from a transcript of an early session in one of the cases I studied (Westerman, 2008). The case involved a female patient in brief therapy with a male therapist. At one point in that session, the patient said that she did not like the way she had been participating in the sessions. The dialogue continued as follows:

> Pt: I just feel like I'm insecure and …(clears throat) we're back to that weakness thing again I guess, but uh, I don't think that I present um…

> Th: (interrupts) (smug smiling) finish that. You don't think you present…

> Pt: uh (dips head down, smiling coyly)…the…any of my strengths. I mean I, uh, cause I get easily flustered and sometimes I get lost with you, um, because I feel pressured and…

>   +Th: What is it…are there things that I'm doing that are contributing to that feeling of pressure?
>
>   Pt: um (looks down, looks at therapist, smiles)…probably (laughs)
>
>   Th: okay (smiling)
>
>   Pt: oh goodness, uh (smiling)
>
>   Th: (interrupts) (smiles) I mean besides things like asking you to finish the sentence or something when you…when you start resisting a thought you had in your mind…are there other things that I…that you can think of that, you know, make it difficult for you…?
>
>   Pt: (sighs) I can't put my finger on it.

Let me begin by placing this exchange in context. Already, even at this early stage of the treatment, the patient had introduced a number of points (like her comment that immediately preceded this excerpt in which she said that she didn't like the way she was acting in the sessions) and then been unable to say much more about those points (as happens in this excerpt). How are we to understand this pattern?

Based on extensive consideration of many sessions from the case and from paper-and-pencil materials that were completed by patient and therapist at multiple points during the treatment, it was possible to determine that the patient was struggling with a conflict. I can characterize her conflict using terms from Benjamin's Structural Analysis of Social Behavior (SASB, Benjamin, 1979): on the one hand, the patient had a wish to be "confirmed as ok as is" (SASB 113) and to be "shown empathic understanding" (114) by the important other people in her life, but she was afraid that if she pursued that wish, other people would "ignore her, pretend she is not there" (126) and "neglect her interests and needs" (125).

We can make sense of the excerpted exchange (and of many, many other interactions that occurred over the course of the treatment) if we think about it in terms of this conflict and the tenets of interpersonal defense theory. To begin with, and again employing SASB terms, the patient could have pursued her wish by "clearly expressing herself" (214) and by "showing herself enthusiastically" (213). She could do this by not only making her initial comments (which, in themselves, were only allusions to things she appeared to want to talk about), but developing them further. However, if she had behaved that way, her fear could have been the result, because the other person (the therapist in the example under consideration) might have responded to her by ignoring/neglecting her. Alternatively, the patient could have protected herself from her feared outcome by foregoing her wish and not showing/expressing herself. She would have been proceeding in that way if she didn't make her allusions at all (and, of course, refrained from elaborating on the allusions she did not make).

Note that if the patient had proceeded in either of these ways, her interpersonal behavior would have been well-coordinated; the elements of her pattern would have meshed together (even though in the latter case, she would have had little to say). The excerpt illustrates how she actually took a very different approach to her interactions with the therapist. Her behavior was characterized by a defensive behavior in which she partially expressed herself (the allusions),

while acting as if she was unable to express herself fully (her recurring difficulty at saying anything more). Therefore, our example exchange offers support for tenet (b) about discourse breaches/coordination failures. To be sure, taken by itself, a single brief excerpt cannot possibly demonstrate that this pattern was a recurring feature of the patient's interpersonal behavior, but it does offer evidence along those lines. In fact, there were a great number of other interactions that were like the one we are considering.

Why did the patient behave this way? According to tenet (c), noncoordinating patterns are attempts to negotiate conflicts between pursuing wished-for relationship outcomes and feared consequences that might occur if a person pursues his or her wished-for outcomes. In the case under consideration, that idea translates into the claim that the patient's behavior represented an attempt to pursue her wish of being confirmed as ok and to be understood empathically, while avoiding being ignored and neglected. Again, the example excerpt provides support. Although it is somewhat difficult to link the patient's behavior in the exchange to being confirmed as ok, it seems reasonably clear that her remark about not liking how she has been acting in the sessions and the way she presented herself as having a hard time elaborating on that remark are attempts to obtain the therapist's empathic understanding.

Furthermore, by partially expressing herself (the allusions) while also acting as if she was unable to express herself fully (her recurring difficulty at saying anything more), she made it quite unlikely that the therapist would respond in the manner she feared (ignoring/neglecting her). In fact, the therapist did not ignore/neglect her at all. In the excerpt and at many other points, he was actively involved with the patient, trying to help her to and/or get her to express herself. Hence, the excerpt also provides support for the part of tenet (d) (the tenet about the feed forward effects of defensive behavior) that maintains that defensive behavior "works" in the sense that it reduces the likelihood that clear-cut instances of feared consequences will occur.

There is much more that can be said about how this single, brief excerpt relates to the tenets of interpersonal defense theory, but the brief remarks I have offered are sufficient to make the point. The example illustrates how in the context of a case study, it is possible to compare observations to multiple tenets of a theory at the same time.

A Possible Objection

It might be argued that my remarks in favor of case studies overstate the case because, even if individual hypothesis-testing studies examine single tenets of a theory, a series of such studies can evaluate the full set of tenets that make up the theory. It certainly is possible for research to proceed in this fashion, and sometimes it does. In fact, taken together, the two lines of hypothesis-testing studies I conducted prior to turning to my current case-study efforts provided support for a good portion of the ideas included in interpersonal defense theory. Nevertheless, it is instructive to consider one way in which this hypothesis-testing approach does not work well in actual practice.

I believe that, all too often, the upshot of pursuing a program of research based on the idea that a set of hypothesis-testing studies can address the different parts of a theory is that the sum of the parts (i.e., the individual studies) ends up equal to far less than the whole (i.e., a

robust theory). Often, when reviewers and editors evaluate individual hypothesis-testing studies, little to no weight is given to the fact that the theoretical tenet examined by a particular study is part of a larger theory and that other studies may have provided support for other parts of the theory. Obviously, in such a situation it does not automatically follow that the new study will/should provide support for the particular tenet under investigation. Whether it will offer support for that tenet is an open empirical question. That's why it makes sense to conduct the new study.

If the results fail to provide support for the tenet, this raises questions about the tenet and also about the theory as a whole (which is not to say that the results of one study count as definitive proof that the tenet and the theory are wrong). But what if the results support the tenet investigated? In this situation, the findings should lead us to have some confidence in the tenet and some additional confidence in the theory as a whole. But often, what actually happens takes a different turn. In many cases, reviewers raise questions about alternative explanations for the study's findings. They raise the possibility that some notion other than the theoretical tenet that motivated the study might account for the findings obtained.

Skeptical suggestions of this sort can play a useful role. A competing explanation might identify the real reason why the investigators obtained their results. Indeed, apparent support from other studies that tested hypotheses related to other tenets of the theory also might be misleading. The whole edifice might be a house of cards.

Keeping these concerns in mind, investigators are likely to devote themselves to follow-up studies aimed at ruling out alternative explanations. Although there is potential merit in proceeding in this way, there also are real dangers, because it is always possible to suggest alternative explanations for any finite set of research results. I believe that progress in the field is often limited by a failure to give sufficient weight to another consideration. Too frequently, investigators, reviewers, and editors do not take into account the following logic: the tenet investigated in the hypothetical study we are imagining is part of a larger theory, we have found support for other parts of that theory, and, therefore, our confidence in the theory should lead us to give some additional weight to the explanation of the findings in question that is based on the theory.

Clearly, we need to be judicious in how we balance the several considerations that are relevant to properly evaluating research results, but if we do not employ a logic that works at the level of the theory as a whole in the way I have just suggested, then research and theory are likely to suffer over time. In my opinion, failing to employ a logic that operates at the level of whole theories has led, in fact, to a situation in many areas of the discipline in which research efforts have been limited to "small" theories that account for highly limited ranges of events. The misguided logic works against developing the kinds of rich theories that we really need if we are going to understand the true complexity of psychological phenomena. Occam's razor is an important methodological principle, but it should be used when it provides a justified way to make an incisive cut in the whole cloth, not to cut away indiscriminately at small pieces of the cloth one at a time.

By contrast to what I have just said about hypothesis-testing studies, case studies have notable advantages when it comes to evaluating a theory as a whole, because investigators employing this research approach compare observations to the multiple tenets of a theory *at the same time*. If repeated observations from a case study *simultaneously* conform to the different parts of a theory, this provides compelling support for the tenets and for the theory as a whole. Observations that simultaneously support the parts of a theory support the theory as a whole for two reasons. The first reason is simple: Such observations support the whole theory because they support its parts. In addition, they support the theory as a whole because they demonstrate empirically that the multiple tenets hang together as a coherent, consistent set of postulates. (Stiles' article does an excellent job of highlighting the issue of coherence/consistency across the multiple tenets of a theory.) They show that one set of observations can meet the requirements of the *set* of predictions that follow from the theory. Moreover, the case-study approach can yield one further kind of support for a theory. Investigators can study other cases and determine whether observations from those cases also support the set of the theory's tenets.

One caveat is in order here. To be sure, it is not always a straightforward matter to determine whether observations in a case study "conform to a theoretical tenet." These determinations involve complex interpretive processes. There is room for error. A researcher might decide that his or her observations support a theoretical tenet when this conclusion is unwarranted. The situation is different in certain important respects in a hypothesis-testing study that employs systematic measurement procedures to define what counts as observations that support a hypothesis. But the issue here is by no means clear-cut. For one thing, as I maintained earlier, interpretation also enters into the quantitative procedures employed in hypothesis-testing studies. A second point brings us back to the earlier discussion about context sensitivity. In case-study research, an investigator might draw the unwarranted conclusion that a certain observation supports a tenet of a theory but, alternatively, he or she also might recognize idiosyncratic ways in which an observation really does support the tenet on occasions when a formal measurement procedure would fail to provide a correct assessment of the behavior in question.

THEORY DEVELOPMENT AS WELL AS THEORY EVALUATION

Stiles' point that case studies can make a significant contribution to the evaluation of theories is a novel idea. In my opinion, it is the single most important idea in the article under consideration. In addition, Stiles maintained that case-study research can contribute in another way as well. Case studies can contribute to theory development by helping us refine, revise, and expand upon a theory of interest.

Stiles (2003, reprinted as Appendix A in Stiles, 2009) illustrated this idea by referring to case studies that led to the elaboration of the APES levels. The case studies I conducted on interpersonal defense theory provide another example of how case studies can be useful for theory development. Prior to conducting the case studies, I had already formulated the tenets of the theory about how defensive interpersonal patterns promote negative outcomes that are distinct from, but related to, feared consequences, and how it also leads to positive outcomes that are distinct from, but also related to, wished-for relationship events. In addition, in collaboration with a colleague, I had already conducted an experimental investigation of those tenets (Dahmen & Westerman, 2007). Nevertheless, the case studies contributed greatly to the previous work on

the two tenets because it helped me better characterize what these typical outcomes of defensive behavior are like and just how they relate to clear-cut examples of feared and wished-for outcomes.

Once again, a caveat is in order. I believe that Stiles was correct when he argued that case studies can contribute to theory building by providing a methodology that is well-suited to evaluating theories and also to developing them, but a tension exists between these two aspects of theory building. When an investigator combines theory justification with theory revision, he or she might commit errors by treating post-hoc explanations of observations (accounts based on the theory as revised in response to those observations) as if they were validated by the observations.

I believe that this is a real danger, but I do not believe that this concern militates against employing Stiles' theory-building approach to case studies in which investigators both evaluate theories and continue to develop them. Rather, this concern calls for great care on the part of case-study researchers. In particular, they must keep in mind that we can only have very limited confidence in a theoretical tenet that makes sense of observations if that tenet was newly revised based on those very observations. But the situation is quite different if a researcher has found that the newly revised tenet also makes sense of further observations from the same case and other cases as well.

The accepted view of hypothesis-testing studies is that they are the proper way to pursue theory validation because they provide a context of justification that is separate from the context of discovery. According to this view, such studies test genuine predictions (i.e., before-the-fact claims about what observations will be made). An investigator who conducted a hypothesis-testing study might offer some speculations in his or her discussion of findings about how the hypotheses examined should be modified, but those post-hoc ideas have a clearly different and secondary status until they themselves are investigated as hypotheses in a subsequent study.

There are merits to proceeding in this fashion but, in my opinion, the issue at hand is quite complicated. In particular, I believe investigators very frequently depart from the "ideal" approach I have just sketched out. For example, within the context of a single hypothesis-testing study, investigators often decide to omit results about one part of a study. Similarly, quite often reviewers or editors suggest omitting material that an investigator included in his or her original submission. Decisions of this sort are typically made when findings from the part of the study in question do not support predictions. Most often when this happens, the investigators, reviewers, and/or editors are convinced by the overall set of findings that the part of the study that was omitted was off track in some way—perhaps because there was something problematic about the methods employed in that part of the study. Proceeding in this manner might be justified, but note that it departs from the logic of keeping the context of justification separate from the context of discovery. Moreover, in actual practice, "mixing" the two research contexts occurs in many other ways as well when researchers conduct hypothesis-testing studies.

These considerations suggest that there is no clear-cut difference between Stiles' model of theory-building case-study research and hypothesis-testing research with regard to keeping justification- and discovery-oriented efforts separate. The point I want to make here is not that

two wrongs make a right. Instead, I believe that it certainly is important to keep in mind the difference between research situations in which observations support genuine predictions as compared to situations in which observations lead to formulating new, clearly tentative ideas, but I also believe that in order to make progress in an even minimally efficient manner it is necessary to carefully mix efforts aimed at justification and discovery.

CONCLUDING COMMENTS

Stiles (2009) has provided us with a major reassessment of what we can learn from case studies. According to him, case studies can make a much greater contribution to research than most psychologists believe, if they are pursued according to Stiles' theory-building approach. One of his claims about the merits of case studies is especially noteworthy: case studies can contribute to theory evaluation.

In the main, I have endorsed Stiles' positive assessment of what case studies can accomplish. In my remarks, I focused on comparisons Stiles drew between theory-building case studies and hypothesis-testing studies. In some respects, I took issue with his claim that case studies are superior to hypothesis-testing investigations with regard to sensitivity to context, because I believe that when hypothesis-testing studies are conducted as explicitly interpretive quantitative inquiries they are context sensitive in certain ways. I argued that the issue about context sensitivity is complicated. I attempted to identify differences that do exist between the two research approaches with respect to sensitivity to context, including specific ways in which case studies are better with regard to this concern.

I also highlighted Stiles' claim that case studies are well-suited to evaluating theories because they provide opportunities to compare observations to the multiple tenets of a theory. I believe this is a very important point. In an attempt to make this technical idea more concrete, I illustrated it with an example from a set of case studies I conducted on the theory of interpersonal defense.

Throughout my comments, I have attempted to point out an array of methodological concerns that should be taken into consideration in evaluating approaches to psychological inquiry. In my opinion, these concerns are complicated and they compete with one another when we try to be responsive to all of them at the same time. All research approaches have a downside. Therefore, well-considered attempts to choose a research method always involve weighing pros and cons of possible ways to proceed. I believe that Stiles' article presents a novel and forceful argument that case-study research actually fares quite well when we weigh the pros and cons in a thoughtful manner. He has shown us that case studies have more to offer than most psychologists realize.

REFERENCES

Benjamin, L. S. (1979). Structural analysis of differentiation failure. *Psychiatry: Journal for the Study of Interpersonal Processes, 42,* 1-23.

Dahmen, B., A., & Westerman, M. A. (2007). Expectations about the long-term consequences of recurring defensive interpersonal behavior. *Journal of Research in Personality, 41,* 1073-1090.

Dewey, J. (1896). The reflex arc concept in psychology. *Psychological Review, 3,* 357-370.

Fishman, D. B. (1999). *The case for pragmatic psychology.* New York: New York University Press.

Heidegger, M. (1962). *Being and time* (M. Macquarrie & E. Robinson, Trans.). New York: Harper & Row.

Keenan, E. O., & Schieffelin, B. B. (1976). Topic as a discourse notion. In C. Li (Ed.), *Subject and topic* (pp. 337-384). New York: Academic Press.

Merleau-Ponty, M. (1962). *Phenomenology of perception* (C. Smith, Trans.). London: Routledge & Kegan Paul.

Osatuke, K., & Stiles, W. B. (in press). Numbers in assimilation research. *Theory & Psychology.*

Stiles, W. B. (1988). Psychotherapy process-outcome correlations may be misleading. *Psychotherapy, 25,* 27-35.

Stiles, W. B. (2003). When is a case study scientific research? *Psychotherapy Bulletin, 38,* 6-11. [reprinted as Appendix A in Stiles, 2009]

Stiles, W. B. (2006). Numbers can be enriching. *New Ideas in Psychology, 24,* 252-262.

Stiles, W. B. (2009). Logical operations in theory-building case studies. *Pragmatic Case Studies in Psychotherapy,* Vol. *5*(3), Article 2, 9-22. Available: http://hdl.rutgers.edu/1782.1/pcsp_journal

Westerman, M. A. (1998). Reconceptualizing defense as a special type of problematic interpersonal behavior pattern: A fundamental breach by an agent-in-a-situation. *The Journal of Mind and Behavior, 19*, 257-302.

Westerman, M. A. (2004). Theory and research on practices, theory and research as practices: Hermeneutics and psychological inquiry. *Journal of Theoretical and Philosophical Psychology, 24,* 123-156.

Westerman, M. A. (2006a). Quantitative research as an interpretive enterprise: The mostly unacknowledged role of interpretation in research efforts and suggestions for explicitly interpretive, quantitative investigations. *New Ideas in Psychology, 24,* 189-211.

Westerman, M. A. (2006b). What counts as "good" quantitative research and what can we say about when to use quantitative and/or qualitative methods? *New Ideas in Psychology, 24,* 263-274.

Westerman, M. A. (2008, September). *Research-informed case study analyses of the theory of interpersonal defense: Findings regarding defensive behavior, countertransference phenomena, and intrapsychic defense mechanisms.* Paper presented at the meetings of the North American Chapter of the Society for Psychotherapy Research, New Haven, Connecticut.

Westerman, M. A. (in press). Conversation analysis and interpretive quantitative research on psychotherapy process and problematic interpersonal behavior. *Theory & Psychology.*

Westerman, M. A., & Foote, J. P. (1995). Patient coordination: Contrasts with other conceptualizations of patients' contribution to the alliance and validity in insight-oriented psychotherapy. *Psychotherapy, 32*, 222-232.

Westerman, M. A., Foote, J. P., & Winston, A. (1995). Change in coordination across phases of psychotherapy and outcome: Two mechanisms for the role played by patients' contribution to the alliance. *Journal of Consulting and Clinical Psychology, 63*, 672-675.

Westerman, M. A., & Prieto, D. M. (2006). Expectations about the short-term functional role
        played by defensive behavior in interpersonal interactions. *Journal of Research in
        Personality, 40,* 1015-1037.
Westerman, M. A., & Steen, E. M. (2007). Going beyond the internal-external dichotomy in
        clinical psychology: The theory of interpersonal defense as an example of a participatory
        model. *Theory & Psychology, 17,* 323-351.
Westerman, M. A., & Steen, E. M. (in press). Revisiting conflict and defense from an
        interpersonal perspective: Using structured role plays to investigate the effects of conflict
        on defensive interpersonal behavior. *Psychoanalytic Psychology.*
Wittgenstein, L. (1958). *Philosophical investigations* (3rd. ed., G.E.M. Anscomb, Trans.). New
        York: Macmillan.