The Question of Verification in Phenomenological Research: Hermeneutic Criteria for Reliability and Validity[1]

Scott D. Churchill
University of Dallas

Abstract: Unlike traditional empirical methods, in which procedures and steps of analysis follow in a linear fashion and can be tracked and thus repeated by others, the empirical methods of the human sciences proceed in a hermeneutical manner, which consists of a "circular" movement that is harder to trace. Not only does the process of research differ, but the nature of the results, as well. A controversy has arisen between "conceptual" and "narrative" modes of thought with respect to psychological description. Conceptual thought yields denotative descriptions which are validated by their correspondence to objective facts; narrative thought yields connotative descriptions (stories) that are validated by their "coherence"--does the story "hang together"? The *Zusammenhang* of experience is indeed a difficult subject for description. Is coherence enough? Without correspondence to some body of facts or observations, on what basis could one assess the reliability of one's methods? Should human science researchers be concerned with such issues? This paper will attempt to outline what is at stake with regard to the questions of reliability and validity in phenomenological research, and will provisionally propose some answers to the questions raised.

When we raise the question of verification in the human sciences, we are asking a question about the experience of truth. This paper is not intended to be an epistemological reflection on the grounds of truth, but rather an empirical reflection on what it means, concretely, to experience a sense of "truth" in the findings communicated to us by human science research.

Whenever I reflect on the question of adequacy of description in the human sciences, I am faced with the initial problem of whether adequacy is to be defined objectively according to some external principle, or whether it is an essentially personal evaluation. The former would obviously make the task of verification easier and would thus be desirable from a methodological viewpoint. However, if the establishment of truth is an essentially personal phenomenon--that is, one which must register in an "authentic" sense in the "mineness" of one's own experience--then we face two problems: first, how is truth initially achieved at the individual level; and second, how is this truth shared.

The sharing of truth is an intersubjective phenomenon that invites a wide range of involvements from open-minded receptivity to critical vigilance. (It is on the latter end of the spectrum what we usually think of verification of science.) Once this sharing has taken place, that is, once a proposition or an interpretation has been verified by a

second observer, on what grounds do we accept this act of verification *as* a verification? If truth is to be experienced authentically, then does it not require that no verification can exist apart from the *experience* of truth itself? The question then becomes, how does a personal verification communicate itself in a compelling way to others? Stated differently, what is the possibility of such verification carrying any weight beyond the experience of the one who actually *does* the verification? Do others have anything more than a blind faith to stand on?

In the first part of this paper I will discuss the nature of verification in terms of "conceptual" and "narrative" approaches to description, and will address the possibility of a convergence between two contrasting criteria for truth ("correspondence" and "coherence"). In the second part I will make a case for a convergence between the traditionally separate issues of validity and reliability.

I

In discussing the question of verification with a colleague last fall at APA, my friend argued that one could verify someone's perception without agreeing with it. That is, he maintained that even as a phenomenologist, one could establish reliability without establishing validity. The example he cited was somebody seeing something on the street that looked like something else that it really wasn't. He would say to this person, "I can see how you could make that mistake."

My immediate response was that one could not make such a statement as a phenomenologist, for to do so would be to

impose one's own perception of reality in such a way as to invalidate the other's perception. As van den Berg (1972) observes in the case of therapy, "Nothing gives us a right to hold our own observation to be truer than the patient's" (p. 46). Would not a disvalidation of another's perception strip the act of verification of all its weight?

When a therapist tells a patient "I can see how you see things the way you do," the patient experiences a sense of relief. I ask, what is the meaning of this relief? Is it not that the patient experiences that his perception of reality has been verified by the therapist, and thus given validity, i.e., truth status? And is this not, in fact, the very foundational premise of phenomenological psychology, namely, that we suspend the mode of belief (and thus the criterion for truth) characteristic of the natural attitude? In the natural attitude, we assume a correspondence between our perception on the one hand, and reality on the other, precisely because in the natural attitude we believe that there is an objective world to which our perceptions may or may not correspond.[2] The phenomenological reduction takes out of play not only this allegedly objective reality, but also any criteria for truth that would naively dictate that there must be a correspondence between an objective world and our experience. Understood phenomenologically, the life-world can be said to exist "in a movement of constant relativity of validity" (Gadamer, 1975, p. 218).

The issue that is raised here can be elaborated by reference to the distinction between conceptual and

narrative modes of thought.[3] The conceptual mode yields descriptions where the emphasis is on reference to facts, and thus statements are "verifiable in the sense that they correspond to an observation . . . validation involves a match or fit between proposition and reality" (Shapiro, 1984). The narrative mode yields descriptions where the emphasis is on sense, often at the expense of exact reference to facts; here, statements are validated to the extent that they are experienced as _making_ sense. Thus, conceptual modes of thinking favor denotative language, which enables a researcher to achieve a more exact _correspondence_ between description and phenomenon; whereas, narrative modes of thinking favor connotative language, which enables one to achieve _coherence_ while articulating less definite correspondences. What we have here are two different criteria for truth: one which sees truth as historical or publically observable facts waiting to be discovered, and the other which sees truth "as the emergent construction of something that makes sense" (Loch, 1977, p. 221). In the latter view, an interpretation "is always a creative act whose historical truth is indeterminate" (Viderman, 1979, paraphrased in Spence, 1982, p. 166).[4] Psychoanalytic therapy, for example, does not require that the analyst's construction of a childhood event be objectively confirmed for it to have "subjective truth value"--i.e., pragmatic value for the present and future of the patient.

To summarize: In the denotative approach of conceptual thought, where a correspondence is sought between facts and

their interpretation, adequacy is judged in terms of accuracy. In the connotative approach of narrative description, the criterion for truth is not a question of whether the description is factually correct, but rather, whether "the story hangs together" (Sarbin, 1984). Adequacy is here judged in terms of coherence or "narrative fit" (Spence, 1982). I submit that what we really need is a hybrid of these two approaches.

It seems to me that coherence without correspondence results in a serious problem for human science researchers, for then we are prone to taking too many liberties with our data. Ricoeur (1977), for example, asks for a narrative intelligibility that we would expect from a story: "A story has to be 'followable', and in this sense 'self-explanatory'" (p. 869). Here, the hanging together of the story is judged solely on the basis of aesthetic experience rather than historical accuracy. To simply ask for a good story leaves one open to too many inadequacies. The problem is that once we move into an aesthetic domain, we become disengaged from an interest in historical or factual truth: "We no longer search for historical accuracy but consider the interpretation in terms of its aesthetic appeal" (Spence, p. 270). But then, using narrative fit alone as a criterion for adequacy, we discover that there are numerous plausible accounts that might never be proven. Thus, "as soon as we admit that, for example, a given account might have a number of different endings, all equally satisfying, we begin to see that establishing narrative fit may be a less definitive outcome

than we might have wished and, as a consequence, a rather shaky basis for making claims about truth value" (Spence, p. 182).

In practice, we need to strive for both coherence and correspondence, and perhaps should consider a third term that embraces both of these criteria. "We see that historical truth, by itself, is not sufficient because the pieces must be fitted into an understandable Gestalt (narrative truth); and, of course, narrative truth could not be maintained if all pieces of the narrative were fabricated (zero historical truth)" (Spence, p. 180). Perhaps the term "connectedness" could serve to suggest the need for a description to be adequately grounded in fact, while expressed in a way that preserves the *Zusammenhang*, or the *Gestalt* of human experience. Rather than using the notion of artistic truth to supplant the concern for factual accuracy, it can be used to amplify and extend our definition of narrative truth. It suggests that not just any narrative will do; that coherence and completeness are necessary but not sufficient; and that an important ingredient of the power to persuade is the aesthetic nature of the narrative" (Spence, p. 270).[5]

At this point I will turn to the issue of reliability and attempt to articulate its connectedness to the issue of validity.

II

"Reliability refers to the consistency of a measuring procedure or instrument. A method of measurement is reliable if it

always produces the same result under the same conditions" (Lewin, 1979). Reliability assumes that one can establish an equivalence of measurement. In the modern sciences, "measurement" has come to mean quantification according to an established standard. When measurement takes the form of a number, then equivalence can readily be established in view of the principle of identity: 10 seconds = 10 seconds. There is, so to speak, a criterion for equivalence that is both formal and objective, and which makes the establishment of reliability a "cut and dried" issue, a matter of "fact". But when the measurement is in the form of a narrative description that uses connotative language, then the equivalence of description is less easy to establish--indeed, not only is the criterion for agreement between two descriptions not defined clearly, but an agreement among those who would judge the equivalence of the two descriptions is difficult to establish. The question "does the story hang together" as a subjective criterion for validity becomes extended to "can these two stories hang together" as a criterion for reliability. What I am suggesting is that there is a Zusammenhang (or "totality," "interconnectedness") of narrative descriptions of a phenomenal reality, just as there is a Zusammenhang of phenomenal reality itself.

Whether one is judging the coherence of a single narrative description, or the cohering of two or more narrative descriptions, the question of coherence remains a matter of personal judgment; reliability thus becomes a second-order question of agreement or "fit" among observations

that are presumed to have achieved a "fit" at the first level.[6] It may be helpful to use a diagram to express the correspondences that I am speaking of here:

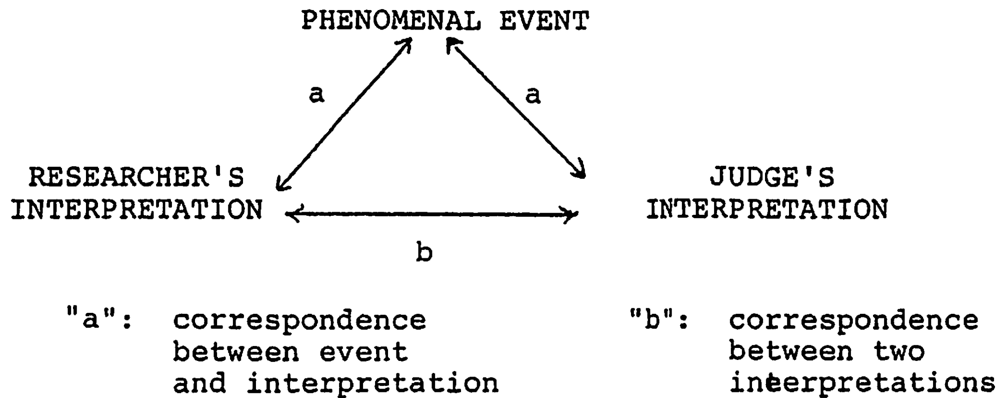

Figure 1

In traditional research, the question of correspondence at the first level (a) can be bracketed in evaluating the correspondence or agreement between measurements (b). For example, one can compare two administrations of an IQ test for their reliability, while begging the question of the validity of the IQ test in the first place. In narrative research, a correspondence of descriptions (b) cannot be established independently of an assessment of the adequacy of fit of descriptions to data (a). Coherence without correspondence results in an arbitrary system of interpretation. The image that comes to mind is a group of "wild analysts" sitting in a parlor room, engaging in blind analysis of a dream. To the extent that the assessment of reliability cannot be made apart from the contexts to which

the descriptions correspond, then reliability becomes inextricaly related to the question of validity, that is, of the correspondence between description and context.

Let us take the example of the interpretation of someone's dream. A comparison of interpretations (for the purpose of determining their adequacy) cannot take place apart from the very context of the interpretations, namely, the patient's life history. In practice, this consists of all that the interpreter knows of the life of the patient: it is, so to speak, the Zusammenhang that is understood only after so many analytic sessions. The part is to be understood from the point of view of the whole. Thus, to compare narrative descriptions of the "part" (the dream) for their equivalence requires a grasp of the totality (the life history) from the perspective of which the descriptions would be made in the first place.[7]

Indeed, one of the problems in establishing reliability of methods in the human sciences is that the researcher is often privy to more data and background information than can be shared with the reader or judge. As Spence (1982) observes in the case of psychoanalytic case analyses, "Because each reader constructs his own 'text' from a mixture of public anecdote and private fantasy, no single set of 'facts' is ever seen by more than one analyst" (p. 236). How can one evaluate the adequacy of one researcher's description to the original data, if the real data are co-constituted within the experience of the researcher? "The concept of the given has here a basically different structure . . . . everything given is created" (Gadamer, 1975, p. 200).

As R.D. Laing (1967) has observed, the subject matter of psychology is not a simple given, or "data," but rather something that is "_taken_ out of a constantly elusive matrix of happenings" (p. 62). Thus, he suggests "we should speak of _capta_ rather than data." Quoting Spence again, "We read a published paper much as we listen to a patient's productions, adding associations to correct gaps in the narrative . . . But because each of us comes to the task with his own set of associations, each of us experiences a slightly different paper" (p. 235). The same must be said when we are reading the data, intuiting a sense of the phenomenal reality being described, and evaluating the adequacy of another's description of that data. Reliability, then, cannot be established by simple observation. Whether two descriptions or measurements agree is a matter of perceiving a coherence bewteen the two descriptions rather than merely a direct correspondence, the latter of which would be simply a close equivalence. The coherence criterion for reliability acknowledges the perspectival nature of qualitative description, and therefore does not presume that two valid descriptions will be equivalent much less the same. Rather, a coherence among descriptions (i.e., among different perceptions of the "same" data) asks only that these varying perspectives be able to "fit" together.[8]

The distinction between validity and reliability becomes blurred in qualitative research precisely to the extent that, in the process of determining whether or not a set of findings can be verified by other researchers, the latter must make their own assessment of how well (if at all) the descriptive findings have reached their target. Let us return to the earlier example where my colleague made the statement, "I can see how you could make that mistake." At face value, this statement affirms the other's perception _as_ a perception, without granting it validity. The issue of reliability can be seen to lie beneath the 'dis'-validation of the other's perception. For, the very fact that I might use my own perception to invalidate another's perception reveals an implicit lack of reliability of perception itself: the judge perceives the situation differently and thus sees the other's perception as a mistake. That is, the assessment of a mistake itself rests on the unreliability of perception.

Thus, the question of validity (i.e., the adequacy of a description with respect to the phenomenon which it characterizes) is already invoked in raising the question of whether or not one can affirm the description. In the case of phenomenological research, it would seem ludicrous to find oneself verifying another's results while at the same time deeming them invalid. The two issues are intimately related insofar as the procedure for assessing the adequacy of the findings is at the same time an assessment of the adequacy of the "measure" used to obtain the

findings (namely, the researcher's perspective on the data). There may certainly be times when one might only partially (or even marginally) agree with other researchers' characterizations of their phenomena, even while being able to affirm their findings in view of the data collected and the perspective adopted in the analysis. Here we would simply speak of a "limited validity" of the findings. In the end, the value of the findings depends on their ability to help others gain insight into the phenomenal reality of human life. "The main function of phenomenological description is to serve as a reliable guide to the listener's own actual or potential experience of the phenomena" (Spiegelberg, 1983, p. 694). Further description from different viewpoints may then supplement, thereby extend, and possibly even radically decenter what remains a forever partial knowledge of human life (Churchill & Wertz, 1985).

<u>Footnotes</u>

1. Paper presented at the Fourth Annual Human Science Research Conference, May 21-25, 1985 at The University of Alberta, Edmonton, Canada.

2. In the psychology of perception, this is known as the constancy hypothsis: for every percept there exists a stimulus event and physidogical event--that is, there is a correspondence between the 'real world' outside and the psychological experience within.

3. Bruner (1984), Loch (1977), Sarbin (1984), Spence (1982).

4. Similarly, Karl Popper argues that "our knowledge claims cannot assert truth or validity as correspondence but only truth as verisimilitude, the appearance of truth." (Shapiro, 1984) Thus, "once it has been realized that close empirical fit [what we have been referring to as the correspondence concept of validity] is no virtue and that it must be relaxed in times of change, then style, elegance of expression, simplicity of presentation, tension of plot and narrative, and seductiveness of content became important features of our knowledge" (Feyerabend, 1975, quoted by Shapiro, 1984).

6. This situation parallels the relationship between what Husserl called the <u>Gegendstand</u> (the actual object of perception) and the <u>noema</u> (the object as "meant"). For Husserl, the Gegendstand would be the ideal sum of all possible noemata, i.e. of all possible persepectives that could be taken on the phenomenon. Given the perspectival nature of perception, any one description is, to use Husserl's language, a noematic apprehension of a transcendent reality. The more noemata embraced by a particular description, the more adequate that description is to the phenomenon. Since no one narrative description can in practice embrace the "whole" phenomenon, then the adequacy of a particular description is judged in view of its limited nature. The judge, too, is capable of only a limited grasp of the phenomenon, and so an equivalence of described observations is either merely coincidental, or is established by the perceived hanging together (or "narrative fit") of the original observation and the judge's own apprehension of the phenomenon.

5. It is important to note here that what we have said so far may only hold up at the idiographic level: in asking of a human event, "how has it happened that it is so" (Gadamer, p.6), the understanding that we come up with may not be generalizable to other circumstances, and thus there is a loss of external validity. "If interpretations are creative rather than veridical and if the [researcher] functions more as a pattern maker than pattern finder, then we may be faced with a glaring absence of general rules." (Spence 292-3)

On the other hand, as we find that general laws "become less persuasive" in the human science, "we are tempted to turn to narrative fit as the guiding criterion for what is true" (Spence 186).

7. To use a mathematical metaphor, we need both numerator and denominator of a numerical relation, in order to establish an equivalence: For example, can the numerators 3 and 5 be equivalent? Only if we know that the denominators are 6 and 10 (3/6 = 5/10) The corespondence of 3 and 5 cannot be established without a grasp of the whole within which they are a part.

Where this analogy breaks down is that in the comparison of descriptions of a human event, there is a <u>common</u> "denominator", namely, the human event itself, in its "totality". Thus we are dealing, in qualititive research, with judgments of equivalence, understood as coherence, where these judgments are necessarily subjective in nature.

8. One problem in speaking of the "fit" between data and description (validity) and the "fit" between two perceptions of the data (the first researcher's description and the judge's own perception, which is the basis for establishing reliability) is that "narrative fit is far easier to establish than narrative failure" (Spence, 183). One can always make a compelling case for one's interpretation, and a judge who agrees with your description can likewise make a convincing case for the "fit" of the original description and his own. If narrative fit is proposed as the criterion for validity and reliability, then we need to be more descriptive about what we mean by "fit".

References

Bruner, J.S. (1984, August). Conceptual and narrative modes of thought. Paper presented at the 92nd Annual Convention of the American Psychological Association, Toronto.

Churchill, S.D. & Wertz, F.J. (1985). An introduction to phenomenological psychology for consumer research: Historical, conceptual, and methodological foundations. In E.C. Hirschman & M.B. Holbrook (eds.), Advances in Consumer Research, Vol. XII.

Feyerabend, P. (1975). Against method.

Gadamer, H.-G. (1975). Truth and method. New York: Crossroad. (Original work published 1960)

Heidegger, M. (1962). Being and time. New York: Harper & Row. (Original work published 1927)

Laing, R.D. (1967). The politics of experience. New York: Ballantine.

Lewin, M. (1979). Understanding psychological research. New York: Wiley.

Loch, W. (1977). Some comments on the subject of psychoanalysis and truth. In J. Smith (ed.), Thought, consciousness and reality, New Haven: Yale University Press.

Ricoeur, P. (1977). The question of proof in Freud's psychoanalytic writings. Journal of the American Psychoanalytic Association, 25, 835-71.

Sarbin, T.R. (1984, August). Prediction and clinical inference: Fourty years later. In R.K. Blashfield (Chair), Clinical versus statistical prediction. Symposium conducted at the 92nd Annual Convention of the American Psychological Association, Toronto.

Shapiro, K.J. (1984, August). Verification: validity or understanding. In R. Alapack (Chair), Humanistic reinterpretation of evidence and related concepts in research. Symposium conducted at the 92nd Annual Convention of the American Psychological Association, Toronto.

Spence, D.P. (1982). Narrative truth and historical truth: Meaning and interpretation in psychoanalysis. New York: Norton.

Spiegelberg, H. (1983). The phenomenological movement (3rd ed.). Boston: Martinus Nijhoff.

van den Berg, J.H. (1972). A different existence: Principles of phenomenological psychopathology. Pittsburgh: Duquesne.

Viderman, S. (1979). The analytic space: Meaning and problems. Psychoanalytic Quarterly, 48, 257-91.