



MEASURING PHENOMENAL ACCURACY: AN EMPIRICAL EVALUATION OF A PHENOMENOLOGICAL METHOD

NEILL WATSON AND MELISSA K. WELCH-ROSS

Department of Psychology, The College of William & Mary,
Williamsburg, Virginia, USA

Forty undergraduates wrote descriptions of personal experiences of self-deception. Procedures adapted from a phenomenological method were used to develop individual situated structural descriptions and a general description under three conditions of participant self-reflection (low, medium or standard, and high). Participants completed the Phenomenal Accuracy Ratings (PAR) for both descriptions. The level of self-reflection used in standard phenomenological method produced the most accurate descriptions. The general descriptions were rated as less accurate than the situated structural descriptions. Results demonstrated the usefulness of the PAR for measuring the phenomenal accuracy of qualitative descriptions of experience. Strategies are proposed for using the PAR to study (a) the commonality vs. individuality in experience and (b) the relation of individual differences to behavior.

The premise that phenomenal experience is a basis of behavior is shared by phenomenological psychology (Giorgi, 1985; Shlien, 1963), client-centered theory (Rogers, 1951, 1959), Adlerian individual psychology (Ansbacher & Ansbacher, 1956), and constructivist theories (Mahoney, 1988; G. J. Neimeyer, 1993; R. A. Neimeyer, 1993) as exemplified in personal construct psychology (Kelly, 1955) and personal narratives (Mair, 1988; R. A. Neimeyer, 1994). These theories conceptualize the person as an agent who actively construes the meaning of events. Thus, a common purpose of these approaches is to describe a person's phenomenal experience.

How well is this purpose accomplished? That is, how well do

Received 23 February 1998; accepted 22 October 1998.

This article is based on a paper presented at the meeting of the Eastern Psychological Association, April 1992, Boston, MA.

We gratefully acknowledge the assistance of Jody Jacobson Alperin in the data collection.

Address correspondence to Neill Watson, The College of William & Mary, Department of Psychology, P.O. Box 8795, Williamsburg, VA 23187-8795, USA. E-mail: npwats@facstaff.wm.edu

the results of a descriptive method correspond to the phenomenal experience? From a constructivist perspective, this question can be analyzed in terms of the epistemological theory that all acts of knowing, whether through everyday personal experience or scientific observation, are constructions by the person or scientist (Kelly, 1955). In the case of a scientific method for describing personal experience, the scientist uses methodological constructs, such as meaning units, to develop a description of the constructs of the experiencing person. Does the method, through its constructive template, yield an accurate or a distorted description of a person's phenomenal experience? Answering this question requires criteria for judging the accuracy of a description of a person's experience. In the present study, we adopted criteria that the research participants used to evaluate descriptions of their experiences developed using a phenomenological method. These criteria of phenomenal accuracy may also be useful in evaluating the results of the methods of other approaches to describing a person's experience.

Describing the essential meaning of an experience is the purpose of a phenomenological method described by Colaizzi (1978) and Giorgi (1985). In this method, the researcher chooses an experience to study—for example, a situation in which one deceived oneself—and identifies his or her own presuppositions about the topic as a basis for questions to elicit descriptions of the experience from participants. The research participants write descriptions of their personal experiences of a specific situation of self-deception. Then, the researcher develops two descriptions of the experience. First, the researcher identifies the meaning units in each participant's description and formulates their meanings in commonsense psychological language. These meanings are synthesized to form a situated structural description of each individual's experience, that is, a description of the fundamental structure of the meaning of an experience of a particular individual. Then the researcher asks the participant, in a clarification procedure, whether the situated structural description reflects the essential meaning of the experience. The researcher incorporates any changes the participant makes into the description. Second, the researcher integrates the common meanings found in the situated structural descriptions of the group of participants to form a general description of the experience. The purpose of the general description is to reflect the essential structure of an experience in terms of commonly shared meanings. Currently, a procedure does not exist for measuring the phenomenal accuracy of these descriptions.

Who is the best judge of phenomenal accuracy? In Giorgi's (1985) phenomenological method, the situated structural description and the general description are referred to other researchers for "confirmation or criticism" (p. 19). In contrast, Colaizzi's (1978) method is to ask the

experiencing person, in a "final validating step" (p. 61), how the description compares to the personal experience. If individuals' experiences are sufficiently unique, it seems to us that the experiencing person is the best judge of how well a description of an experience corresponds to the experience itself. The uniqueness of an individual's perceptions is emphasized in the individuality corollary of Kelly's (1955) theory of personal constructs: "Persons differ from each other in their construction of events" (vol. 1, p. 55). Also, Shlien (1963), in his account of Rogers's (1951, 1959) phenomenological approach to personality, implies that the experiencing person is the best judge of how well a description reflects phenomenal experience: "Each person is unique. No one else can ever completely know his experience The closest approach to another's experience is to see it through his own eyes, insofar as possible" (p. 320). Rogers (1951) implemented this assumption in client-centered therapy by looking to the client for indications that the therapist's empathy was accurate.

In the present study, we asked research participants to judge the phenomenal accuracy of their own situated structural descriptions and of the general description. Phenomenal accuracy was defined as how well, in the opinion of the experiencing person, a description corresponded to that person's experience. We developed an instrument for obtaining Phenomenal Accuracy Ratings (PAR) that consisted of three scales for judging (a) the essential meaningfulness of the description, (b) the omission of essential constituents, and (c) the inclusion of irrelevant constituents. The distinction we made between essential and irrelevant constituents is similar to the distinction that Giorgi (1988) made between defining and characteristic constituents in his account of Husserl's (1962/1977) discussion of personal meaning. We expected that the ratings of essential meaningfulness would correlate negatively with the ratings of omissions and with the ratings of irrelevancies, thereby showing the interrelationships among these criteria of phenomenal accuracy. Further, we expected that the PAR would have predictive value in (a) determining the degree to which the situated structural description and the general description corresponded to each individual's personal experience and (b) examining the relative effectiveness of particular variations on phenomenological method.

More specifically, with the PAR we evaluated the accuracy of descriptions developed using phenomenological method by comparing them with descriptions developed using a control procedure and with descriptions developed using a procedure designed to enhance phenomenal accuracy. In phenomenological method, the researcher's preconceptions about the experience under investigation can bias the results and thus render them inaccurate. Standard phenomenologi-

cal method employs a clarification procedure to reduce this threat to accuracy: The research participant is asked whether the description reflects the meaning of the experience and the researcher incorporates changes indicated by the participant (Colaizzi, 1978). This procedure employs the self-reflection of the participant as a check on the potentially invalidating effect of the researcher's preconceptions.

Alternatively, a modification of this standard method could enhance the accuracy of the situated structural description by employing participant self-reflection at an additional point in the procedure. In the standard method, the researcher identifies the meaning units essential to the situated structural description. In the modified method, the participant would identify the meaning units. Specifically, the participant would be asked to identify the meaning units that are essential to the experience under study. This procedure could reduce the effects of any biases stemming from the researcher's own preconceptions that the clarification procedure with the participant fails to reduce.

In the present study we hypothesized that participant self-reflection in the development of the situated structural description would increase the accuracy of that description and, consequently, of the general description. To test the hypothesis, standard phenomenological method was compared with a procedure that employed less self-reflection, as well as with the procedure just described, which increased the level of self-reflection. By increasing the extent to which the participant was involved in developing the situated structural description, we expected to reduce the omission of essential constituents and the inclusion of irrelevant constituents, thereby increasing accuracy ratings of the situated structural description and, therefore, of the general description. However, we also expected that, across all levels of participant self-reflection, the general description would be rated less accurate than the situated structural descriptions because of the uniqueness of individual experiencing.

METHOD

Participants

Twenty-seven female and 23 male undergraduates volunteered for a study described as one for which they must be able to recount an experience of self-deception. Twenty-four females and 16 males completed the study. Participants received credit for fulfilling a research participation requirement in an introductory psychology course.

Materials

A complex experience consisting of several constituents may be more likely than an experience involving very few constituents to show differences in accuracy ratings as a function of the level of participant self-reflection. Therefore, participants in the present study were asked to provide personal accounts of the experience of self-deception, which was found in previous research (Fischer, 1985) to be characterized by a number of essential constituents.

Experience of Self-Deception

The following directions, obtained from Fischer's (1985) study of self-deception, were given to participants.

I am interested in learning more about what happens when people deceive themselves. Would you be willing to write a description of a situation in which you deceived yourself? Further, would you please include in your description a characterization of what you did in your effort to deceive yourself, e.g. how did you think about the situation, what did you tell yourself and others about it, etc.? Finally, would you also describe how you came to realize that you had been deceiving yourself? (Fischer, 1985, modified by the present authors)

Phenomenal Accuracy Ratings (PAR)

The participant rated the situated structural description and the general description to indicate the degree to which the description (a) reflected the essential meaning of the experience ("The description reflects the essential meaning of my experience of self-deception to me."), (b) omitted essential constituents of the experience ("The description includes all important aspects of my experience."), and (c) included irrelevant constituents ("The description adds aspects that were not part of my experience."). The ratings were on a scale from 1 ("not at all") to 7 ("very much"). In scoring, the scaling was reversed for omission of essential constituents so that the score would reflect the degree of omission directly rather than inversely.

Design and Procedure

Participants signed up for one of three time slots, which were subsequently designated as a low (control), medium (standard), or high

(modified) self-reflection condition. The medium self-reflection condition was the standard phenomenological method (Colaizzi, 1978; Giorgi, 1985) previously described. The standard method provided the participant with an opportunity to clarify the situated structural description developed by the researcher. The low self-reflection (control) condition eliminated the clarification procedure. In the high self-reflection (modified) condition, a procedure was changed in the standard method: The participant, instead of the researcher, identified the meaning units in the development of the situated structural description.

The study was administered by a female graduate student who had been given specialized training in phenomenological analysis so that she remained blind to the hypotheses. Participants attended three sessions, which were several weeks apart. The first session consisted of a group administration in which all participants were given one hour to write a description of an experience of self-deception (see previous section, "Experience of Self-Deception"). The first session was identical for the medium (standard) and low (control) self-reflection groups. However, the high self-reflection (modified) group was asked before the end of the first session to identify the meaning units in their descriptions according to the following instructions:

Now please reread your description and write what you think are the most important parts of your experience of being self-deceived, as indicated in your description. This includes the parts that, if they were absent, would lead to the experience not being one of self-deception. In other words, these are to be the themes that are important and *essential* to your experience being labeled "self-deception."

Situated Structural Descriptions

Between the first and second sessions, the researcher developed a situated structural description for each participant. As described in the preceding section, each participant in the high self-reflection group had identified the meaning units in the original description. Following the same instructions, the researcher identified the meaning units in each of the original descriptions of the other two groups after reading the entire description to grasp its meaning as a whole. In the next step, she formulated the meaning units in commonsense psychological language. Finally, she synthesized the formulated meaning units into a situated structural description for each participant.

Each participant returned individually to a second session, read the situated structural description together with the participant's original description of the experience, and rated the situated structural description on the PAR. Then, in a clarification procedure, participants

in the high self-reflection (modified) group and the medium self-reflection (standard) group were given an opportunity to change the situated structural description so that it better reflected their experience. After clarification, these participants again rated the description on the PAR. Participants in the low self-reflection (control) group were not allowed to clarify the description.

General Descriptions

Between the second and third sessions, the researcher developed three general descriptions, one for each of the three self-reflection groups. The criterion for the general description was that it contain all the formulated meaning units that were common, at least implicitly, to the situated structural descriptions. A separate description was developed for each group in order to see whether the predicted differences among the three groups in the accuracy of the situated structural descriptions also would be reflected, as predicted, in the general descriptions. Finally, each participant returned individually to a third session, read the general description together with the participant's original description of the experience, and rated the general description on the PAR.

RESULTS

Situated Structural Descriptions: Examples

An example of a situated structural description from each of the three groups is presented below. For the medium and high self-reflection groups, the examples illustrate additions (in italics) made by many participants during the clarification procedure. These additions and other changes were the basis for improved accuracy ratings after the clarification procedure and for the better accuracy ratings by the medium and high self-reflection groups as compared to the low self-reflection group (see subsequent section "Analyses of Accuracy Ratings").

The reader can compare the example situated structural description with the general description for each group (see the next section), in order to see common constituents and omitted and irrelevant constituents in the general description, as indicated by participants in their accuracy ratings (see subsequent section "Analyses of Accuracy Ratings").

Low self-reflection (control) group

Self-deception is convincing oneself of something which is not true, even though one is subconsciously aware of the truth. One tells

others one's false belief. One realizes one has been deceiving oneself when someone else tells the truth.

Medium self-reflection (standard) group

Self-deception is ignoring one's instincts and reality. It involves ignoring what others say which confirms the truth. One deceives oneself because one does not like reality. One does not tell others of one's self-deception. It requires much strength to realize one has been deceiving oneself. *Feelings of depression would tie in with not understanding or accepting the truth and reality.*

High self-reflection (modified) group

Egocentricity causes one to deceive oneself. Self-deception is creating a false image of oneself which one would like others to believe [and which] *the person believes about him/herself. The person's needs or wants may cause self-deception in order for attainment.* In this way, self-deception involves deceiving others and may become unpleasant if others discover this.

General Descriptions

The general descriptions of the experience of self-deception developed by the researcher for each of the three groups are presented below.

Low self-reflection (control) group

Self-deception involves denying the truth and convincing oneself of something which is not true. One tells others excuses which confirm one's false belief. One is subconsciously aware of the truth, but one acts in accordance with the false belief because one does not like the truth. One realizes one has been deceiving oneself when one accepts the truth.

Medium self-reflection (standard) group

Self-deception is motivated by one's desire for something. It occurs when one does not like reality. Self-deception involves denying the truth and holding a false belief. One lies to oneself. One also has a distorted perception of reality. One tells others the false belief and makes excuses for one's behavior. One realizes one has been deceiving oneself when one accepts the truth.

High self-reflection (modified) group

Self-deception is motivated by one's desire for something. It involves denying the truth and holding a false belief. One acts according to one's false belief and may make excuses to justify one's actions. One realizes one has been deceiving oneself when one accepts the truth.

Analyses of Phenomenal Accuracy Ratings

Intercorrelations among the three accuracy ratings were computed. For the situated structural descriptions, Pearson correlations showed a strong negative relation between meaningfulness and omission of essential constituents, $r(38) = -.56$, $p < .001$, and between meaningfulness and the inclusion of irrelevant constituents, $r(38) = -.65$, $p < .001$. A positive correlation was found between the omission of essential constituents and the inclusion of irrelevant constituents, $r(38) = .85$, $p < .001$. A multiple regression analysis showed that 41% of the variance in the meaningfulness ratings of the situated structural descriptions was explained by inclusion of irrelevant constituents; the omission of essential constituents did not account for any unique variance.

For the general description, meaningfulness was negatively correlated with the omission of essential constituents, $r(38) = -.64$, $p < .001$, and with the inclusion of irrelevant constituents, $r(38) = -.65$, $p < .001$. Omission of essential constituents and inclusion of irrelevant constituents were positively correlated, $r(38) = .60$, $p < .001$. A multiple regression analysis showed that 41% of the variance in the meaningfulness ratings of the general description was explained by the inclusion of irrelevant constituents and that an additional 9% was explained by the omission of essential constituents. Thus, the three ratings that comprise the Phenomenal Accuracy Ratings converged as expected.

A 3 (Self-reflection: low, medium, high) \times 2 (Gender) \times 2 (Description: situated structural, general) MANOVA, with repeated measures on the description variable, was performed on the three accuracy ratings. This analysis tested the hypotheses that (a) higher levels of participant self-reflection would increase the accuracy of the situated structural description and the general description, and (b) the general description would be less accurate than the situated structural description across all levels of self-reflection.

An alpha level of .05 was used for statistical tests. The multivariate analysis was used to control for a Type I error in subsequent multiple univariate tests, because the dependent variables, as expected theoretically, were shown to be intercorrelated by Bartlett's test of sphericity, $\chi^2(3, N = 40) = 44.72$, $p < .001$. The assumptions of homogeneity and normality were met for the between-subjects variances on the multivariate test, Box's $M = 57.60$, $F(21, 612) = 1.37$, $p < .05$. However, for univariate tests, the assumption of homogeneity for between-subject variances was not met on ratings of omission of essential constituents in the situated structural description, and on ratings of inclusion of irrelevant constituents in both the situated structural and the general descriptions, Cochran's $Cs(6, 6) > 0.47$, $ps < .05$. Therefore, nonparametric

univariate tests were used to pursue between-subject effects on these ratings. Parametric tests were used on the other ratings, which had acceptable, similarly skewed, distributions as well as homogeneous variances across groups. Also, because the variances and covariances were heterogeneous for the within-subjects variables, Mauchly's $W = 0.35$, $\chi^2(5, N = 40) = 34.01$, $p < .001$, Greenhouse-Geisser's epsilon (0.59) was used to adjust the degrees of freedom for a more conservative F -test to analyze within-subject effects in univariate tests subsequent to significant multivariate effects.

Table 1 shows cell means and standard deviations for males and females combined, because main effects and interactions involving gender were nonsignificant in the MANOVA. In this and the other analyses, it is possible that the absence of effects for gender was due to the small cell sizes for each gender (see footnotes to Table 1) and the consequent lack of power for detecting gender differences.

TABLE 1 Phenomenal Accuracy Ratings as a Function of Level of Self-Reflection and Type of Description

Type of description/Variable	Level of self-reflection		
	Low ^a (control)	Medium ^b (standard)	High ^c (modified)
Situated structural			
Meaningfulness			
<i>M</i>	6.19	6.91	6.54
<i>SD</i>	0.66	0.30	0.52
Omission of essential constituents			
<i>M</i>	2.44	1.00	1.15
<i>SD</i>	1.32	0.00	0.56
Inclusion of irrelevant constituents			
<i>M</i>	2.25	1.00	1.31
<i>SD</i>	1.07	0.00	0.86
General			
Meaningfulness			
<i>M</i>	5.25	6.00	5.77
<i>SD</i>	1.13	1.27	1.24
Omission of essential constituents			
<i>M</i>	3.06	1.82	2.46
<i>SD</i>	1.65	1.17	1.85
Inclusion of irrelevant constituents			
<i>M</i>	2.56	2.27	2.31
<i>SD</i>	1.60	1.19	2.06

^a $n = 16$.

^b $n = 11$.

^c $n = 13$.

In the MANOVA, a main effect was found for self-reflection, Wilks's $\lambda(6, 64) = 0.67, p < .05$. A univariate three-way ANOVA and subsequent pairwise comparisons of means with Tukey's HSD tests were performed on ratings of essential meaning. Results showed that in comparison to the low self-reflection group, the medium group rated both the situated structural and the general descriptions as more meaningful, $F(2, 34) = 3.58, MSE = 0.95, p < .05$. Kruskal-Wallis one-way ANOVAs and pairwise comparisons of means with Mann-Whitney *U*-tests (both tests corrected for ties) were performed on ratings of omission of essential constituents and on ratings of inclusion of irrelevant constituents. Results showed that in comparison to the low group, the medium and high groups rated the situated structural description as omitting fewer essential constituents, $\chi^2(2, N = 40) = 20.23, p < .0001$, and as including fewer irrelevant constituents, $\chi^2(2, N = 40) = 17.48, p < .001$. As predicted, the high and medium self-reflection groups reported greater phenomenal accuracy than the low group. Contrary to prediction, the high group did not differ from the medium group.

The MANOVA also revealed a main effect for description, Wilks's $\lambda(3, 32) = 0.56, p < .001$. Therefore, univariate three-way ANOVAs were performed on the three accuracy ratings. The results showed that in comparison to the situated structural description, the general description was rated by all groups as less meaningful, $F(1, 20) = 23.90, p < .01$, as omitting more essential constituents, $F(1, 20) = 11.34, p < .01$, and as including more irrelevant constituents, $F(1, 20) = 7.56, p < .05$. As predicted, the general description showed lower phenomenal accuracy than the situated structural description.

A comparison of the variances of the accuracy ratings of the situated structural description and the general description across all groups showed greater variance for the general description on essential meaningfulness, $t(38) = 5.00, p < .001$, omission of essential constituents, $t(38) = 2.60, p < .05$, and inclusion of irrelevant constituents, $t(38) = 2.92, p < .01$. The greater dispersion of ratings of the general description included a higher percentage of participants whose ratings were beyond the scale midpoint to the endpoint that represented lowest accuracy: on essential meaningfulness, 7.5% for the general description versus 0.0% for the situated structural description; on omission of essential constituents, 17.5% versus 2.5%; and on inclusion of irrelevant constituents, 12.5% versus 2.5%. Thus, participants showed less agreement about the accuracy of the general description than about the situated structural descriptions, as well as lower accuracy ratings of the general description.

To examine the effect of the clarification procedure, univariate

analyses were performed on each of the three accuracy ratings to compare ratings of the situated structural description before and after clarification. The medium and high self-reflection groups in these analyses were larger than in the analysis reported earlier because less attrition had occurred by this point in the study. The clarification procedure was not used with the low self-reflection group. Although the dependent variables were shown to be intercorrelated by Bartlett's test of sphericity, $\chi^2(3, N = 30) = 10.66, p = .01$, a multivariate analysis was not done to control for a Type I error in the multiple univariate tests because the data were not amenable to Box's *M*-test for homogeneity of dispersion matrices. Therefore, the probability of a Type I error was controlled by adopting an alpha of .01 as the significance level for each of the three univariate tests. Nonparametric tests were used because the within-subject variances were heterogeneous on all three ratings, $ts(28) > 2.39, ps < .05$.

In examining the clarification procedure, Wilcoxon Matched-Pairs Signed-Ranks tests on the three accuracy ratings showed that, in comparison to the description before clarification, participants tended to rate the description after clarification as more meaningful, $z = -2.37, p < .05$ (Mean Ranks = 0.0 and 4.0, before and after, respectively), and participants rated the description after clarification as omitting fewer essential constituents, $z = -2.52, p = .01$ (Mean Ranks = 4.5 and 0.0, before and after, respectively). Therefore, ratings of phenomenal accuracy were higher after the clarification procedure.

To examine differences between the high and medium self-reflection groups with this larger sample, a 2 (Group: medium, high) \times 2 (Gender) \times 2 (Clarification: before, after) ANOVA, with repeated measures on the clarification variable, was performed on the ratings of essential meaning. This analysis was not conducted on the other two accuracy ratings because they did not have homogeneous variances, Cochran's $Cs(7, 4) > 0.56, ps < .05$. On the ratings of essential meaning, a main effect for group, $F(1, 26) = 9.03, MSE = 0.44, p < .01$, showed that in comparison to the high group, the medium group rated the situated structural description as more meaningful both before ($Ms = 6.64$ and 6.06 , medium and high groups, respectively) and after ($Ms = 6.93$ and 6.50 , medium and high groups, respectively) clarification. Mann-Whitney *U*-tests showed that the medium group tended to rate the situated structural description before clarification as including fewer irrelevant constituents than did the high group, $p < .05$ (Mean Ranks = 13.00 and 17.69, respectively). Contrary to prediction, the medium self-reflection group reported greater accuracy of the situated structural description than did the high self-reflection group.

DISCUSSION

In the present study we asked research participants to judge the accuracy of descriptions of their personal experiences that were developed using a phenomenological method. More specifically, participant judgments on the Phenomenal Accuracy Ratings were used (a) to determine the degree to which the situated structural description and the general description corresponded to each individual's personal experience and (b) to examine the relative effectiveness of particular variations on phenomenological method.

The Phenomenal Accuracy Ratings were an internally consistent measure with predictive utility. The intercorrelations among the three ratings showed that the essential meaningfulness of the description of an experience was inversely related to the omission of essential constituents and to the inclusion of irrelevant constituents. Also, the omission of essential constituents and the inclusion of irrelevant constituents each accounted for unique variance in the ratings of essential meaningfulness of the general description of the experience of deception that was developed in the present study. These results suggest the importance of including both of these variables in a measure of meaningfulness.

We interpret the pattern of findings as supporting the phenomenal accuracy of the descriptions developed using the level of participant self-reflection in the phenomenological method described by Giorgi (1985) and Colaizzi (1978). The standard method, which included the clarification procedure, produced better results on all three participant accuracy ratings than did the control method, which did not include a clarification of researcher-identified constituents. Specifically, for the situated structural descriptions, the standard method resulted in fewer omissions of constituents that participants considered essential to their experiences, as well as the inclusion of fewer constituents that participants considered irrelevant. Also, the standard method resulted in higher ratings of essential meaningfulness for both the situated structural description and the general description.

An alternative interpretation of the higher phenomenal accuracy ratings for the standard method as compared with the control method is that participants in the standard method were more involved in the study than were the control participants. That is, participants in the standard method were allowed to clarify the situated structural description, whereas participants in the control method were not. However, contrary to this alternative interpretation, the participants in the high self-reflection method, who were the most highly involved in

that they identified the meaning units in addition to clarifying the situated structural description, judged the description as lower in phenomenal accuracy than did the participants in the standard method. Therefore, a more likely explanation is that the difference in accuracy ratings between the standard and control methods is a function of the clarification procedure rather than a function of greater involvement per se.

Contrary to our hypothesis, the results indicate to us that increasing participant self-reflection, by having participants identify the meaning units of their experiences in the procedure for developing the situated structural descriptions, does not improve upon the results that can be obtained using standard phenomenological method. In comparison to the standard group, the high self-reflection group rated the situated structural description both before and after the clarification procedure as less meaningful. Also, in comparison to the standard group, the high self-reflection group tended to rate the situated structural description before clarification as having more irrelevant constituents. Thus, according to the participant's own judgment, the researcher was better than the participant at identifying the essential meanings of the participant's experience without including irrelevancies.

However, when the participant clarified the situated structural description after the researcher had integrated the essential meanings, the accuracy of the description improved. In comparison to the group that did not use the clarification procedure, the groups that used it reported that the description omitted fewer essential constituents and included fewer irrelevant constituents. Also, the groups that used the clarification procedure rated the description after clarification, in comparison to before clarification, as omitting fewer essential constituents and tended to rate it as more meaningful. Thus, the clarification procedure improved phenomenal accuracy, whereas developing a situated structural description on the basis of participant-identified meaning units did not. In summary, the accuracy of the situated structural description appears to be greatest when the procedures of standard phenomenological method are used: Essential meanings are identified and integrated into a description by the researcher and the description is clarified by the participant.

As predicted, participants judged the general description as less accurate than the situated structural description on all three ratings. Mean ratings of the general description fell between the midpoints and the endpoints of the scales, suggesting to us that participants, as a group, judged the general description to be more accurate than not. Nevertheless, the development of a general description from the situated structural descriptions resulted in mean accuracy ratings that were

lower than the accuracy ratings obtained for the situated structural descriptions. Differences between the two descriptions in the variances of all three ratings showed that participants agreed less about the accuracy of the general description than about the accuracy of the situated structural description. Also, in comparison to the situated structural description, more participants rated the general description between the midpoints and the endpoints that represented "not at all" accurate on all three ratings. Thus, participants disagreed about the accuracy of the general description, with some participants rating it low in accuracy.

One possible explanation for this result is that the researcher's training in phenomenological method was specifically limited. We provided the researcher with specialized training in the method so that she would be blind to the hypotheses. Perhaps a more thoroughly trained phenomenologist could develop a general description that participants judge to be as accurate as the situated structural descriptions. However, we think that the lack of agreement concerning the accuracy of the general description suggests that individual differences in personal experiences may be great enough to preclude the development of a single general description that all persons would rate high in accuracy.

The disagreement among participants about the accuracy of the general description has implications for the issue of the commonality of human experience. To what extent are there commonly shared meanings in personal experience? In personal construct theory, the issue can be conceptualized in terms of Kelly's (1955) individuality corollary, "Persons differ from each other in their construction of events" (vol. 1, p. 55), and his commonality corollary, "To the extent that one person employs a construction of experience which is similar to that employed by another, his psychological processes are similar to those of the other person" (vol. 1, p. 90). According to these propositions, an individual's personal experience has constituents that are unique to that individual and may also have constituents that are shared with other individuals. The extent to which individuals share constituents of experience is the question of commonality. In the present study participants disagreed about the essential meaningfulness of the general description, about the relevance of its constituents, and about the omission of constituents. This disagreement calls into question the commonality of meanings related to the specific experience of self-deception, as well as the commonality of meanings more generally.

If researchers use phenomenological method to study commonalities in experience, then additional research is needed that uses measures, such as the Phenomenal Accuracy Ratings, to evaluate whether

constituents of a general description of an experience are indeed common to the entire group that is studied. High ratings of essential meaningfulness and low ratings of inclusion of irrelevancies, together with high agreement among participants on both ratings, would be necessary to conclude that all the constituent meanings of the description are common to the participants' experiences. High ratings of omission of essential constituents, together with high agreement among participants, would indicate that one or more common meanings are missing. Using these criteria, successive approximations of a final general description would be drafted by the researcher and judged by the participants, and specific irrelevancies and omissions would be identified and corrected, insofar as possible, until ratings of the description are optimized. This method would allow one to determine whether a general description can be developed that all participants will judge as accurate.

In the present study, the method for measuring phenomenal accuracy made the research participant the judge of the accuracy of the situated structural description and the general description. In contrast to our approach, Giorgi (1986) assumed that the best judge of the accuracy of a description is someone expert in phenomenological method. As discussed earlier, we think the present findings suggest the uniqueness of meanings in individuals' experiences. If individuals' experiences are unique, then the assumption that one individual, including an expert phenomenologist, is the best judge of the accuracy of a description of another individual's personal experience seems questionable to us.

In addition to studying commonalities, phenomenological method can be used to obtain a better understanding of individual differences in experience and their relation to behavior. The procedure we used to measure phenomenal accuracy can be used to fine-tune the study of individual differences, and we expect this procedure would lead to the development of self-report measures with high criterion-related validity. More specifically, adaptations of phenomenological method can be used to develop situated structural descriptions of a phenomenal experience, for example, thoughts or images that precede or accompany a behavior such as restrained eating. The Phenomenal Accuracy Ratings can serve as criteria for determining whether or not satisfactory descriptions have been achieved. Then, the constituents of the descriptions can be used as items in a self-report measure which is subjected to tests of criterion-related validity, using the behavior as the criterion.

A measure of phenomenal accuracy is, in our view, a measure of the validity of the results of any method for describing personal experience. Phenomenal validity is distinct from other types of validity used in psychology, for example, criterion-related validity as defined

by Anastasi (1976). Giorgi (1988) also made this distinction, interpreting Husserl's (1962/1977) discussion of personal meaning as a basis for defining validity in phenomenological method. We suggest that phenomenal validity is a prerequisite for criterion-related validity. That is, self-report measures that are high in phenomenal validity should have higher criterion-related validity than self-report measures that are low in phenomenal validity.

The present study provides a model for using quantitative methods to measure the validity of qualitative data. The procedure can be used with a variety of descriptive approaches other than phenomenological method, including participant observation, ethnography, narrative analysis, and protocol analysis (Campbell, 1996; Denzin & Lincoln, 1994; Ericsson & Simon, 1993), Adlerian life style assessment (Shulman & Mosak, 1988), and the repertory grid technique in personal construct psychology (Kelly, 1955). Specifically, the present study illustrates how researchers can assess the accuracy of descriptive data and suggests ways that researchers may be able to improve accuracy through integrating qualitative and quantitative methods. In this respect, our approach addresses concerns that descriptive methods may not adequately reflect the meanings in a person's experience (e.g., R. A. Neimeyer, 1994).

Some researchers may object to integrating qualitative and quantitative methods on philosophical grounds and therefore reject the notion that quantitative methods can be employed to achieve better qualitative data (for a related discussion see Kvale, 1989). We suggest that quantitative methods can be used to assess validity while maintaining, as a central concern, the importance of assuring that the researcher understands, as much as possible, the individual's personal experience. Measures of phenomenal validity can be applied both when the goal is to understand a particular individual's experience and when the goal is to understand aspects of experience that are common to everyone in a group. In both cases, the process of obtaining phenomenally valid descriptions of experience involves a dialectical relation between the experiencing person and the researcher (for a related discussion see Taylor, 1990). In this dialectical process, measures of phenomenal accuracy are an important gauge for evaluating whether or not qualitative procedures, such as phenomenological method, have achieved their purpose of describing personal experience accurately.

REFERENCES

- Anastasi, A. (1976). *Psychological testing* (4th ed.). New York: Macmillan.
- Ansbacher, H. L., & Ansbacher, R. R. (Eds.). (1956). *The individual psychology of Alfred Adler*. New York: Harper & Row.

- Campbell, D. T. (1996). Can we overcome worldview incommensurability/relativity in trying to understand the other? In R. Jessor & A. Colby (Eds.), *Ethnography and human development: Context and meaning in social inquiry* (pp. 153-172). Chicago: University of Chicago Press.
- Colaizzi, P. F. (1978). Psychological research as the phenomenologist views it. In R. S. Valle & M. King (Eds.), *Existential-phenomenological alternatives for psychology* (pp. 48-71). New York: Oxford University Press.
- Denzin, N. K., & Lincoln, Y. S. (Eds.) (1994). *Handbook of qualitative research*. Thousand Oaks, CA: Sage Publications.
- Ericsson, R. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data* (Rev. ed.). Cambridge, MA: MIT Press.
- Fischer, W. F. (1985). Self-deception. In A. Giorgi (Ed.), *Phenomenology and psychological research* (pp. 118-154). Pittsburgh, PA: Duquesne University Press.
- Giorgi, A. (1985). A sketch of a psychological phenomenological method. In A. Giorgi (Ed.), *Phenomenology and psychological research* (pp. 8-22). Pittsburgh, PA: Duquesne University Press.
- Giorgi, A. (1986). Theoretical justification for the use of descriptions in psychological research. In P. D. Ashworth, A. Giorgi, & A. J. J. de Koning (Eds.), *Qualitative research in psychology: Proceedings of the International Association for Qualitative Research in the Social Sciences* (pp. 3-22). Pittsburgh, PA: Duquesne University Press.
- Giorgi, A. (1988). Validity and reliability from a phenomenological perspective. In W. J. Baker, L. P. Mos, H. V. Rappard, & H. J. Stam (Eds.), *Recent trends in theoretical psychology* (pp. 167-176). New York: Springer-Verlag.
- Husserl, E. (1977). *Phenomenological psychology* (J. Scanlon, Trans.). The Hague, Netherlands: Martinus Nijhoff. (Original work published 1962)
- Kelly, G. A. (1955). *The psychology of personal constructs* (Vols. 1-2). New York: Norton.
- Kvale, S. (Ed.) (1989). *Issues of validity in qualitative research*. Lund, Sweden: Studentlitteratur.
- Mahoney, M. J. (1988). Constructive metatheory: I. Basic features and historical foundations. *International Journal of Personal Construct Psychology*, 1, 1-35.
- Mair, M. (1988). Psychology as storytelling. *International Journal of Personal Construct Psychology*, 1, 125-137.
- Neimeyer, G. J. (1993). *Constructivist assessment: A casebook*. Newbury Park, CA: Sage.
- Neimeyer, R. A. (1993). An appraisal of constructivist psychotherapies. *Journal of Consulting and Clinical Psychology*, 61, 221-234.
- Neimeyer, R. A. (1994). The role of client-generated narratives in psychotherapy. *Journal of Constructivist Psychology*, 7, 229-242.
- Rogers, C. R. (1951). *Client-centered therapy: Its current practice, implications, and theory*. Boston: Houghton-Mifflin.
- Rogers, C. R. (1959). A theory of therapy, personality, and interpersonal relationships, as developed in the client-centered framework. In S. Koch (Ed.), *Psychology: A study of a science: Vol. 3. Formulations of the person and the social context* (pp. 184-256). New York: McGraw-Hill.
- Shlien, J. M. (1963). Phenomenology and personality. In J. M. Wepman & R. W. Heine (Eds.), *Concepts of personality* (pp. 291-330). Chicago: Aldine.
- Shulman, B. H., & Mosak, H. H. (1988). *Manual for life style assessment*. Muncie, IN: Accelerated Development.
- Taylor, D. S. (1990). Making the most of your matrices: Hermeneutics, statistics, and the repertory grid. *International Journal of Personal Construct Psychology*, 3, 105-119.

Copyright of Journal of Constructivist Psychology is the property of Psychology Press (UK) and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.