Journal of Personality and Social Psychology
1993, Vol. 64, No. 3, 356-364

0022-3514/93/$3.00

Inferring Interpersonal Traits From Behavior: Act Prototypicality Versus Conceptual Similarity of Trait Concepts

Rainer Riemann and Alois Angleitner

This article investigates 2 models of the cognitive process underlying trait ratings: The trait-to-trait process states that trait inferences are guided by the conceptual similarity among traits. We presented Ss with 6 fictitious persons. Each was described by acts referring to 1 of 6 traits. Ss rated the targets on interpersonal trait terms. If trait ratings are gathered immediately after the presentation of behavioral information about a single target (Study 1) these ratings correspond closely to the prototypicality ratings of the acts on the trait concepts (*behavior-to-trait process*). However, if Ss have to keep the behavioral information about several targets in memory (Study 2), the trait ratings are guided by the conceptual relations among the concepts under study (*trait-to-trait process*).

The contrast between mechanical collection of data in science and the reliance on poorly understood human judgments in psychological assessment have made many psychologists skeptical about the dependability of the latter kind of data. Among the many claims and empirical demonstrations of shortcomings of social judgments (see Nisbett & Ross, 1980), the so-called *systematic distortion hypothesis* (D'Andrade, 1965, 1974; Shweder, 1975, 1982; Shweder & D'Andrade, 1979, 1980) gained a prominent position because it questioned the validity of traditional personality assessment. Shweder and D'Andrade's (1979, 1980; D'Andrade, 1965, 1974; Shweder, 1982) central tenet is that

> *inferences* about personality contain a systematic bias in that propositions about "what is like what" are substituted for propositions about what is likely, and *memory* for personality relevant events contains a systematic bias in that attitudes, affects, and behaviors that are conceptually associated (e.g., "aggression" and "dominance," "disagrees" and "criticizes") are recalled as if they covaried. (Shweder, 1982, p. 66)

The arguments against the traditional (trait) theory of personality put forward by Shweder and D'Andrade and several other researchers (e.g., Berman & Kenny, 1976; Berman, Read, & Kenny, 1983; Bourne, 1977; Chapman & Chapman, 1967, 1969; Mischel, 1968) have been countered by a number of authors for several reasons (Block, 1977; Block, Weiss, & Thorne, 1979; Borkenau, 1986b; Borkenau & Ostendorf, 1987; DeSoto, Hamilton, & Taylor, 1985; Epstein, 1979; Jackson, Chan, & Stricker, 1979; Lamiell, Foss, & Cavenee, 1980; Romer & Revelle, 1984; Weiss, 1979; Weiss & Mendelsohn, 1986). A central point in the discussion of the systematic distortion hypothesis is the question how should behaviors shown by a person be coded for scientific analysis? The focus of the present article, however, is the question how do judges derive trait inferences on the basis of limited behavioral information?

D'Andrade (1965, 1974) and Shweder (1982; Shweder & D'Andrade, 1979, 1980) have demonstrated a close relation between the structure of rated behavior matrices (i.e., correlation matrices computed from memory-based ratings across a subject sample) and conceptual association matrices (i.e., similarity of meaning judgments among the same concepts used for the ratings) and concluded that "conceptual proximity is used as a guide for estimating cooccurrence probability" (Shweder, 1982, p. 95). Romer and Revelle (1984), Borkenau (1986b), and Borkenau and Ostendorf (1987) argued that multiple semantic relationships between acts and trait-descriptive terms are an integral part of trait inferences.

This semantic relationship is defined operationally by means of subjects' ratings of how good an example an act description is of a trait category (*prototypicality ratings;* e.g., Cantor & Mischel 1977, 1979a; Mervis & Rosch, 1981; Rosch, 1975). It has been shown that behaviors, nominated as members of a trait category, vary in their prototypicality for the respective trait (Buss & Craik, 1980, 1983, 1984). For example, the act "Mary set the goals for a group" is judged as highly prototypical for dominance, and "Mary insisted on doing the driving on the trip" is a low prototypical exemplar. The analysis of *multiple prototypicality ratings* (i.e., the prototypicality of each act is rated in reference to several trait concepts; Angleitner & Demtröder, 1988; Borkenau, 1986b; Borkenau & Ostendorf, 1987; Buss, 1985; Buss & Craik, 1986) indicates that acts are closely related to more than one trait category. An act that yields high prototypicality ratings for *gregarious* is likely also to be judged as a good example of *outgoing.*

Borkenau (1986b) and Borkenau and Ostendorf (1987) have found that the structure of on-line recorded behavior frequencies, if they are weighted by multiple prototypicality ratings, corresponds closely to the structure of memory-based ratings.

Rainer Riemann and Alois Angleitner, Department of Psychology, Universität Bielefeld, Bielefeld, Federal Republic of Germany.

The research reported in this article was supported by a grant from the University of Bielefeld (OZ 2766).

We thank Peter Borkenau, Willem K. B. Hofstee, Robert A. Wicklund, Bogdan Wojciszke, and three anonymous reviewers for their helpful comments on drafts of this article.

Correspondence concerning this article should be addressed to Rainer Riemann, Abteilung Psychologie, Universität Bielefeld, Postfach 8640, D-4800 Bielefeld 1, Federal Republic of Germany.

However, this relation is of approximately the same size as the structural relationship between memory-based ratings and conceptual association matrices. Thus, it remains unclear whether judges who provide memory-based ratings recur to their knowledge about conceptual relations among the concepts or rely on the multiple implications of the available behavioral information.

Given the close correspondence between correlations among weighted on-line recorded behavior frequencies and conceptual associations (e.g., Borkenau & Ostendorf, 1987), correlations among proximity matrices are not a sensitive tool to answer this question. The aim of the present article is to test more directly which of these sources of information judges use for trait inferences. To disentangle the effects of both kinds of information, we presented homogeneous act descriptions to our subjects who then rated the targets' standing on a set of related trait concepts not directly covered by the act descriptions.

Processes Underlying Trait Inferences

Both positions outlined here do not explicate the cognitive process that leads from observed behavior to trait inferences. We compare two models of this process: The first stems from the arguments presented by Shweder and D'Andrade (1979, 1980; D'Andrade, 1965, 1974; Shweder, 1982). It claims that judges use behavioral information to infer not more than one trait and then infer other traits from their conceptual similarity to the trait already inferred (*trait-to-trait process*). The second model is based on the studies by Romer and Revelle (1984), Borkenau (1986b), and Borkenau & Ostendorf (1987). It incorporates the finding that most behaviors have implications for many traits, so that less behavior-relevant traits may also be inferred from the behavioral information in addition to the more appropriate traits (*behavior-to-trait process*).

To elaborate these models, we refer to Srull and Wyer's (1989) model of person memory and judgment, which provides a good account of the trait-to-trait process. Consider that a rater is informed about several of a target person's behaviors. Each act is subject to on-line processing at the time when it is observed. The on-line processing results in long-term memory representations of behaviors and inferences (e.g., elaborations and additions). Many of these inferences are interpretations of the target's behaviors in terms of more abstract trait concepts (e.g., Hastie & Park, 1986; Moskowitz & Roman, 1992; Park, 1989; Srull & Wyer, 1989; Winter & Uleman, 1984). Trait–behavior clusters are formed when multiple behaviors are encoded in terms of a trait concept. Partially on the basis of the on-line trait inferences, subjects try to form a general evaluative concept of the person (Srull & Wyer, 1989, p. 59).

When asked to provide memory-based trait ratings subjects retrieve the stored information on the target. If a representation is found that has direct implications for the characteristic in question, judgments are based on the implications of this concept without reviewing individual behaviors. If such a representation cannot be found, subjects will recur to the general evaluative concept of the person.

However, the rules governing this decision are not specified within Srull and Wyer's (1989) model. If raters apply strict criteria (e.g., *gregariousness* does not equal *dominance*), their judgment will be based on the evaluation-based person representation. Alternatively, if less strict rules are used at this point (e.g., both concepts refer to the person's interpersonal behavior), the ratings may be based on knowledge structures relating traits to each other (trait-to-trait process). These have been described as implicit personality theories (e.g., Bruner & Tagiuri, 1954; Rosenberg & Sedlak, 1972; Schneider, 1973).

Note that it is important to postulate that judges in a multiple trait rating task retrieve trait information on a target only once as long as it carries any implications for the multiple ratings and retrieve the same trait again after they have encountered unrelated concepts. Otherwise, this model would not explain the close correspondence between correlations among memory-based ratings and ratings of conceptual similarity.

The behavior-to-trait process postulates that judges relate behaviors to multiple concepts. This may not be accomplished by an automatic processing of behavioral information (see Bargh, 1984) or the on-line processing of this information with an impression formation task in mind (see Hastie & Park, 1986; Hastie & Pennington, 1989; Smith, 1989). To base their memory-based judgments on the multiple implications of behavioral information, judges must retrieve a descriptively unbiased sample of this information from memory. However, this may be the case only if judges do not engage in forming an impression of the target and focus their attention on the meaning of the behaviors instead. We expect that the efficiency of this process (i.e., the correspondence between memory-based ratings and the multiple prototypicality of the behaviors) will be affected by the delay between the encoding of behaviors and the judgment and the complexity of the rating task (e.g., the number of acts exhibited by the targets or the number of targets observed simultaneously).

Overview of the Present Studies

In two experiments we studied trait inferences under different task demands. We described hypothetical target persons by sets of acts that had originally been generated by subjects as examples for trait categories. Thus, they represented naturally chunked units (Cohen, 1981a; Newtson, 1976) that perceivers had extracted from the observed stream of behavior.

In our first experiment, we presented act descriptions to subjects rated as either high or low in prototypicality for the trait category for which they had been nominated. Subjects provided trait ratings immediately after reading each act list. Under these conditions, we expected the behavior-to-trait process to operate. We assumed that low prototypical act descriptions hamper the impression formation process and the formation of trait–behavior clusters. By definition, low prototypical acts are only loosely related to the trait concept, and subjects may hardly detect a common theme among a list of low prototypical act descriptions. Thus, conceptual similarity should poorly correspond to the trait ratings based on low prototypical act descriptions.

We designed the second experiment to examine task demands that we expected to favor the trait-to-trait process. Subjects had to learn act descriptions of six target persons before the trait ratings. In this experiment, we varied the number of acts in each description of a target, the interval between the

presentation of the act descriptions and the ratings, as well as the prototypicality status of the acts. We expected that the high memory load (induced by presenting six sets of act descriptions before the ratings) would result in a decrease of the availability of behavioral information.

Our model of the trait-to-trait process and the distortion hypothesis do not specify which concept will initially be chosen by a judge to anchor subsequent ratings (see Weiss & Mendelsohn, 1986). However, the correspondence between trait ratings and conceptual similarity data depends largely on the identification of the concept where judges anchor their ratings. We used two strategies to select an anchor in our experiments: First, we considered the trait concept for which the acts had been nominated. This seemed to be the most appropriate a priori choice, because we used relatively homogeneous act descriptions for each target that had all been nominated for a single trait. Second, the trait for which a target received the highest ratings was selected. The simple reasoning behind this choice is that subjects anchor their judgments at the concept that most appropriately summarizes the target's behaviors.

Study 1

Method

Overview and Design

Two groups of subjects were given six lists of act descriptions each characterizing a hypothetical target person. Each list consisted of 10 act descriptions. All acts within a list had been nominated for 1 trait concept, but different traits were selected for each of the targets. One group was given descriptions that were highly prototypical for the traits, and the other group received low prototypical descriptions. Subjects were asked to rate the targets on 35 traits or to indicate that they had no information about the trait in question. We analyzed the ratings on 23 interpersonal traits (Wiggins, 1979) in this set. To compare the two models of trait inferences, multiple prototypicality ratings for each behavior and judgments of the conceptual similarity among the traits were obtained independently.

Subjects and Procedure

Multiple prototypicality ratings of the acts. The prototypicality ratings were provided by 20 paid students, 10 men and 10 women. Six American act lists (Buss, 1981), each consisting of 100 acts, were translated into German. The acts had been nominated for the traits *dominant, submissive, gregarious, aloof, agreeable,* and *quarrelsome.* Half of each gender rated acts with a male or female actor, respectively (e.g., "*He* monopolized the conversation." vs. "*She* monopolized the conversation."). In six separate sessions, each of the 600 acts was judged for its prototypicality status on each of the six trait categories (multiple prototypicality rating). The acts were printed on cards. The serial order of the trait categories and acts was randomized for each subject. The instructions for this rating were taken from Buss and Craik (1980), who adapted the procedure of Rosch and Mervis (1975). Subjects judged on a 7-point scale how good an example of the trait category the various acts were. They indicated their ratings by sorting the cards into respective ballot boxes on which labels of the scale points were printed. These ranged from *very good example of the category* to *fits the category very poorly.* The other boxes were used to indicate intermediate judgments. The agreement among the raters for the prototypicality ratings was good. Internal consistency estimates ranged from .84 (for dominance) to .93 (for gregariousness, quarrelsomeness, and agreeableness). For cross-cultural comparisons of the German data with data gathered in the United States, see Angleitner, Buss, and Demtröder (1990).

For each of the six traits, the 10 acts that yielded the highest and the 10 acts that yielded the lowest prototypicality ratings were presented to a sample of 12 judges (6 women and 6 men) who provided prototypicality estimates of each act for each of the 23 interpersonal traits in our list. The trait adjectives were *agreeable, aloof, ambitious, arrogant, calculating, coldhearted, conciliatory, distant, distrustful, dominant, gregarious, ingenuous, kind, lazy, open-minded, quarrelsome, reserved, self-assured, sociable, submissive, tactful, unassuming,* and *warm.*[1] The procedure to collect these ratings was the same as for the initial prototypicality ratings. However, subjects were provided a bipolar scale ranging from +3 (*very good example*) to −3 (*strongly disconfirming example*). The agreement between raters was .94 (averaged across all traits) and ranged from .84 (ingenuous) to .97 (kind).

Conceptual similarity between the trait adjectives. A sample of 12 paid students (6 men and 6 women) judged the semantic similarity among the 23 trait adjectives. They rated the similarity of all possible pairings of the traits using a 7-point rating scale (1 = *very similar,* 4 = *neutral,* and 7 = *opposite*). The traits *agreeable* and *gregarious,* as well as *dominant* and *quarrelsome,* were judged to be quite similar, whereas the adjective pairs *submissive–dominant, quarrelsome–agreeable,* and *gregarious–aloof* were seen as quite opposite in their semantic meanings. The agreement between the raters was high. The average correlation was .65, and Cronbach's alpha was .95.

Trait ratings. Two samples of 20 paid students of the University of Bielefeld (each comprising 10 men and 10 women) completed the trait inference task. The students were recruited by written announcements posted at several places in the university.

We compiled two sets of hypothetical act descriptions of target persons for each of the six traits by selecting either the 10 most prototypical or the 10 least prototypical act descriptions for the respective trait on the basis of the initial multiple prototypicality ratings. The descriptions were presented to half of the subjects in a female or male version, together with a corresponding first name. For example, the act descriptions "Mary set the goals for a group" and "Mary made decisions without consulting the others involved in them" were part of the high-prototypical act descriptions of the dominant target person, Mary.

The act descriptions of the six target persons were given in written form to the subjects. One half of the subjects was given high prototypical, the other half was given low prototypical act descriptions. Subjects were instructed to imagine that they had spent a month together with each of the six target persons and that, during this time, the target had exhibited the respective behaviors. Before presentation of the first act list, subjects were informed that they had to provide multiple trait ratings of the targets. After reading each description, subjects judged on the following page the degree to which the targets' behaviors could be described and summarized by each of 35 traits. A 5-point scale was provided for each rating (1 = *the trait is very descriptive of the target* and 5 = *the trait is not at all descriptive of the target*). As an alternative to the rating, subjects were given the opportunity to indicate that they had no information referring to the trait in question.

Dependent Variables

The main dependent variables in our studies were measures of divergence between the trait ratings and the independently obtained mul-

[1] The German terms were *freundlich, reserviert, ehrgeizig, arrogant, berechnend, kühl, verträglich, unnahbar, mißtrauisch, dominant, gesellig, naiv, gefällig, faul, aufgeschlossen, streitsüchtig, zurückhaltend, selbstsicher, kontaktfreudig, unterwürfig, taktvoll, anspruchslos,* and *warm.*

tiple prototypicality and conceptual similarity ratings. To measure the extent to which a subject's trait ratings for a target reflected the multiple prototypicality of the act descriptions first we calculated separately for each target the mean prototypicality of the 10 acts for each of the 23 traits (averaged across 12 judges who provided the multiple prototypicality ratings). Bivariate regression analyses were then performed to predict each subject's trait ratings for each target from the mean prototypicality ratings. The mean of the absolute values of the residuals (the difference between estimated trait ratings and observed trait ratings) was the measure of the correspondence between trait ratings and multiple prototypicality ratings. Note that higher values indicate less correspondence. Finally, because we did not have any hypothesis regarding differences between trait concepts, the means of the absolute values of the residuals were averaged across traits.

Thus, for example, a deviation score of 0.60 indicated that a subject's actual ratings on the 23 trait terms, which were provided on a 5-point scale, deviated 0.6 points on average from ratings predicted from the mean prototypicality ratings. Although these deviation scores could not be interpreted without considering the variance of each subject's ratings, a within-subjects comparison was appropriate.

Measures for the trait-to-trait process were obtained for two anchor points: the trait for which the act descriptions had been nominated and the trait on which the target received the highest rating. The composite judgments of the conceptual similarity between these categories and the remaining 22 traits were entered into the bivariate regression analysis (predicting each subject's trait ratings) to compute residuals and proceed as described earlier. The three measures are referred to further on as *deviation from multiple prototypicality* and *deviation from conceptual similarity* anchored at *corresponding trait* and *highest trait rating*.

Results

No reliable differences between male and female targets were found. Thus, in the following analysis the data for both sexes were combined.

Table 1 lists the mean deviation measures for the behavior-to-trait and trait-to-trait processes. An analysis of variance (ANOVA) comprising the between-subjects factor prototypicality status of the act descriptions (high or low) and the within-subjects factor deviation measure (deviation from multiple prototypicality and deviation from conceptual similarity anchored at *corresponding trait* or *highest trait rating*) revealed a main effect for deviation measures, $F(2, 76) = 31.15$, $p < .001$, and a significant interaction of both factors, $F(2, 76) = 13.00$, $p < .001$. Although the respective means indicate that ratings based on high prototypical act descriptions were predicted more accurately by all measures, this effect was not statistically significant, $F(1, 38) < 1.0$.

After comparing the means across both groups, we found that trait ratings were more precisely predicted by the acts' multiple prototypicality status than by the conceptual similarity ratings anchored at the corresponding trait, $t(39) = 6.69$, $p < .001$, or at the highest rating, $t(39) = 5.21$, $p < .001$. Separate analyses for high and low prototypical act descriptions revealed that the deviation from multiple prototypicality was significantly smaller than the deviation from conceptual similarity anchored at the corresponding trait or highest rating within both groups; $t(19) = 3.15$, $p < .01$, and $t(19) = 4.00$, $p < .001$, for the high prototypical act descriptions; $t(19) = 7.33$, $p < .001$, and $t(19) = 3.44$, $p < .01$, for low prototypical act descriptions. There was no significant effect for the comparison of low versus high prototypical act descriptions for any of the deviation measures.

Discussion of Study 1

Consistent with our hypothesis, the trait ratings in this study reflect the multiple prototypicality of the acts pretty well. Our measure of the behavior-to-trait process is a better predictor of trait ratings than the two a priori anchored measures of the trait-to-trait process. Thus, the results lend support to our view that trait ratings are generated according to the behavior-to-trait process under the conditions of this experiment. The absence of reliable differences between the two groups presented with high versus low prototypical act descriptions indicates that behavior-to-trait processing is not just a default option, if the impression formation process and the formation of trait–behavior clusters are hampered by low prototypical act descriptions, which bear only little relevance to the trait concepts under study.

Study 2

The trait rating task in Study 1 was designed to meet the conditions that were hypothesized to be favorable for the behavior-to-trait process. The accessibility of a target's behaviors was high, and there was no interference with descriptions of other targets. In Study 2 we presented the descriptions of all targets before collecting the trait ratings. This sequence of presentation and trait ratings should result in a higher memory load and therefore reduce the accessibility of the behavioral informa-

Table 1
Mean Deviation Measures as a Function of Prototypicality Status of Act Descriptions (Study 1)

Deviation measure	High prototypicality ($n = 20$)	Low prototypicality ($n = 20$)	Combined
Multiple prototypicality	0.54_a	0.57_a	0.56_a
Conceptual similarity anchored at			
Corresponding trait	0.57_b	0.64_b	0.60_b
Highest rating	0.59_c	0.60_c	0.59_b

Note. Subscripts refer to within-subject t tests comparing the deviation measures within each column. Means sharing a common subscript within a column do not differ significantly at the $p < .05$ level.

tion. However, to preserve the reliability of the trait ratings, care was taken that the subjects learned which behaviors were presumably performed by which target person.

In addition, we manipulated experimental conditions that may affect the processing of the behavioral information. We varied the interval between the presentation of the act descriptions and the trait ratings (immediately after the presentation of all act descriptions vs. 15-min delay), the prototypicality status (high vs. low) of the act descriptions, and the number of acts (three vs. seven) in each description of a target person.

A delay between the presentation of the targets and the trait ratings should further reduce the accessibility of the single acts. In several studies (Borkenau, 1986a; Cantor & Mischel, 1979b; Cohen, 1981b) it has been shown that subjects' memory for person-related information is less accurate the longer the delay between presentation and recall. Although recall of information that is inconsistent with a preexisting schema is worse than recall of consistent information, this distortion does not increase with the length of the retention interval (see Borkenau, 1990; Cooper, 1981). We hypothesize that the delay hampers behavior-to-trait processing but has little effect if inferences are based on conceptual similarity.

In Study 1 we have found a behavior-to-trait processing even for high prototypical act descriptions. We manipulated the act descriptions' prototypicality status in our second study to explore whether under conditions that otherwise favor trait-to-trait processing there is a shift to behavior-to-trait processing associated with the presentation of low prototypical act descriptions. We expect that the formation of trait–behavior clusters is more difficult if the behaviors are less related to a trait concept.

Finally, we varied the demand on memory during the learning process by presenting three versus seven act descriptions per target. Rothbart, Fulero, Jensen, Howard, and Birrell (1978) have shown that higher memory load has an effect on the processing (representation in memory) of social information. We expect that the higher memory load induced by the presentation of seven acts per target should negatively affect the behavior-to-trait process.

Method

Subjects

Eighty-nine students participated in this study. Depending on the time subjects needed to complete the experiment, they were paid between 4 DM and 11 DM. Subjects were randomly assigned to one of the eight experimental groups. Nine subjects failed to learn the association between targets' names and behaviors within the time limit and were excluded from the experiment. Eight of these subjects had been assigned to the group presented with seven low prototypical act descriptions per target, the remaining subject was assigned to the group presented with three low prototypical act descriptions. Half of the remaining 80 subjects were women.

Procedure

Descriptions of targets. Each act description of male targets used in Study 1 was split randomly into two sets of act descriptions consisting of three or seven acts, respectively. This resulted in 24 hypothetical descriptions of persons: The acts had been nominated for one of six traits, they had either a high or a low prototypicality status for this trait, and they consisted of three or seven single acts.

Learning. Subjects were run individually. They were placed in front of a microcomputer and first read the following instruction:

> Imagine that you have spent a month together with six different target persons who performed several observable behaviors during this time. These behaviors will be presented to you on the computer screen together with the name of the person who performed the behaviors. It is your task to learn these behaviors and to form an impression of the target person. Afterward the behaviors will be presented to you again, and then you will have to name the person who performed each behavior.

The presentation of the act descriptions and the learning procedure were controlled by the computer. All act descriptions of a target were presented simultaneously on the screen together with the target's name. Subjects were allowed as much time as they liked to read each description. After pressing a key, subjects were presented the description of the next target until they had inspected all six descriptions. Then the behaviors were presented individually on the screen together with the names of all six target persons. By pressing marked buttons on the keyboard, subjects indicated their responses. They were given immediate feedback on their choice. If the answer was wrong, they were given a second choice. If the second response was also wrong, the right name was shown on the screen. After subjects had completed the whole set of acts, the number of wrong first responses was calculated. The learning phase was ended if this number was less than four, otherwise the whole set was presented again. Subjects who did not reach the criterion within approximately 1 hr were excluded from the experiment.

Trait ratings. Trait ratings of each target were gathered as described in Study 1. However, subjects in this experiment were not given the opportunity to indicate that they had no information referring to the trait in question.

One half of the subjects completed the trait ratings immediately after they had finished the learning phase successfully (immediate condition). The other subjects worked at filler tasks for 15 min that were unrelated to the present experiment (word fluency and preferences for stimulating situations).

Semantic similarity and prototypicality ratings. The similarity data and the prototypicality ratings for male targets of Study 1 were used in this experiment.

Analysis

The procedure to derive the dependent measures was the same as in Study 1. However, some targets yielded equally high mean ratings on more than one trait. Thus, the highest rating was not considered as an anchor point in this study because it could not be identified unequivocally.

Results

The number of trials, which subjects passed in learning the association between targets' names and the act descriptions, varied significantly between groups. The mean number of trials were 1.80 in the group presented with three high prototypical act descriptions per target, 2.15 (three low prototypical acts), 1.55 (seven high prototypical acts), and 5.95 (seven low prototypical acts). An ANOVA comprising the between-subjects factors prototypicality status of acts and number of acts per target yielded significant main effects for both prototypi-

cality status, $F(1, 76) = 25.85$, $p < .001$; number of acts, $F(1, 76) = 14.40$, $p < .001$, and a significant interaction $F(1, 76) = 18.79$, $p < .001$. It was especially difficult to learn the association between targets' names and the seven low prototypical act descriptions.

Table 2 shows that the deviation from the conceptual similarity ($M = 0.75$) anchored at the corresponding trait yielded lower means than did deviation from multiple prototypicality ($M = 0.72$) under nearly all conditions. Only under one condition (seven high prototypical act descriptions, no delay) did the mean for the deviation from conceptual similarity equal that of the multiple prototypicality measure. We conducted a Prototypicality × Number of Acts × Delay analysis of covariance (ANCOVA) on the deviation measures (repeated measures factor), adjusting for number of trials. Note that this adjustment had no effect on the repeated measures factor and its interactions. The ANCOVA revealed a significant main effect for deviation measure, $F(1, 72) = 35.16$, $p < .001$; a Deviation Measure × Prototypicality interaction, $F(1, 72) = 8.54$, $p < .01$; a Deviation Measure × Number of Acts interaction, $F(1, 72) = 8.36$, $p < .01$; and a marginally significant triple interaction of deviation measure, prototypicality, and number of acts, $F(1, 72) = 3.36$, $p < .075$. The remaining effects were not significant (all $ps > .10$). The ANCOVA results indicate that a Prototypicality × Number of Acts interaction, which was significant in the corresponding ANOVA, $F(1, 72) = 5.70$, $p < .05$, can be attributed to the higher number of learning trials under the seven low prototypical acts condition.

Mean deviation from multiple prototypicality was significantly higher than mean deviation from conceptual similarity. An inspection of the respective marginals showed that the means for the deviation from conceptual similarity at corresponding trait were nearly equal for low and high prototypical act descriptions ($M = 0.73$ vs. $M = 0.72$) and for descriptions composed of three or seven acts ($M = 0.73$ vs. $M = 0.71$).

The means for deviation from multiple prototypicality were 0.73 versus 0.78 (high vs. low prototypical descriptions) and 0.78 versus 0.73 (three vs. seven acts). Thus, there was no evidence that the presentation of low prototypical act descriptions or the presentation of a small number of behaviors was detrimental to the trait-to-trait processing in this experiment.

A closer examination of the marginally significant Deviation Measure × Prototypicality × Number of Acts interaction revealed that deviation from multiple prototypicality was highest when three low prototypical acts were presented. Under this condition the maximal difference between the deviation measures (combined across delay) was observed (deviation from multiple prototypicality $M = 0.85$; deviation from conceptual similarity $M = 0.77$); $t(39) = 4.60$, $p < .001$. The smallest difference between deviation from multiple prototypicality ($M = 0.76$) and deviation from conceptual similarity ($M = 0.75$) was observed under the seven high prototypical acts condition, $t(39) = 1.39$, *ns*. These results contradict our hypotheses that predict the behavior-to-trait processing most likely to occur under the three low prototypical acts condition and a trait-to-trait processing under the seven high prototypical acts condition.

Discussion of Study 2

The results of this study provide evidence that under the conditions of Study 2 subjects infer traits without reviewing all behaviors. Our measures of the trait-to-trait process turned out to be a superior predictor of trait ratings if the behavioral information about all six targets was presented before the rating task. This result was observed even when the trait concept for which the descriptions had been nominated was chosen as an anchor to derive a measure for the trait-to-trait process. This anchoring resulted in a marginally poorer correspondence between conceptual similarity ratings and trait ratings in Study 1 than the anchoring at the highest trait rating.

We hypothesized that the presentation of low prototypical act descriptions hampers trait-to-trait processing. However, the measure of this process tended to be less affected by the presentation of low prototypical act descriptions than the measure of the behavior-to-trait process. There is no evidence that subjects derived trait ratings by a behavior-to-trait processing under any condition of this experiment. This explains the absence of any interpretable effects for the variation of the number of acts

Table 2
Mean Deviation Measures as a Function of the Prototypicality Status of Act Descriptions, Number of Acts in Each Description, and Delay Between Presentation of Act Descriptions and Trait Ratings

	High prototypicality				Low prototypicality				
	3 acts		7 acts		3 acts		7 acts		
Deviation measure	No delay	Delay	No delay	Delay	No delay	Delay	No delay	Delay	Combined
Multiple prototypicality	0.69_a	0.72_a	0.72_a	0.80_a	0.76_a	0.95_a	0.72_a	0.69_a	0.75_a
Conceptual similarity achored at corresponding trait	0.68_a	0.68_b	0.72_a	0.77_a	0.69_b	0.85_b	0.68_a	0.68_a	0.72_b

Note. $n = 10$. Subscripts refer to within-subject t tests comparing the deviation measures within each column. Means sharing a common subscript within a column do not differ significantly at the $p < .05$ level.

within each description and the manipulation of the delay between the presentation of information and the collection of trait ratings. Both manipulations should affect behavior-to-trait processing but not trait-to-trait processing.

It is important to note that the experimental procedure in Study 2 was not immoderately biased to trigger the trait-to-trait process. It was guaranteed that each subject paid intensive attention to the behaviors performed by a target and was able to identify which behaviors were associated with whom for the vast majority of the acts. A side effect of this procedure was that the number of learning trials (i.e., the number of times subjects were presented with the complete set of act descriptions) differed significantly between groups. This effect was controlled statistically in our analyses. Repeated exposure to the act descriptions may result in a higher accessibility of the acts as a basis for the ratings. The respective marginals, however, provide no support for this interpretation.

General Discussion

The important progress to be made in the study of trait inferences is to understand the rules that link behavioral information to trait judgments instead of deciding whether trait ratings are distorted. The central concern of this research is to specify the conditions under which trait ratings are guided by the conceptual similarity between traits or reflect the conceptual relations between behaviors and trait categories. From the results obtained in our experiments we conclude that the process by which subjects derive trait inferences from act descriptions depends mainly on task demands: The experimental procedures of Study 1 result in a behavior-to-trait processing; those of Study 2 favor a trait-to-trait process.

There were several important procedural differences between the two studies that might have produced the different outcomes: (a) In Study 1 subjects provided trait ratings immediately after the presentation of act descriptions for a single target person. Subjects in Study 2 were given the behavioral information about several targets before the rating task. (b) Subjects were expecting multiple trait ratings in Study 1, whereas in Study 2 an impression formation instruction was given. (c) Subjects also had repeated training on the association between acts and targets in Study 2 but not in Study 1.

The order of exposure to the stimulus material and judgments most likely has an effect on the accessibility of behavioral information at the time when the ratings are provided. The reduced accessibility under the conditions of Study 2 should result in less correspondence between multiple prototypicality ratings and trait ratings. Without additional assumptions, however, this effect does not explain the change from behavior-to-trait processing to trait-to-trait processing. The variation of instructions (multiple trait ratings vs. impression formation) may have an effect on the on-line processing of the behavioral information at encoding and the representation of information in memory. If they expect multiple trait ratings, subjects may be motivated to draw multiple inferences from the behaviors on-line or create relatively direct memory representations of this information. Under an impression formation instruction subjects may encode behaviors in terms of a single trait concept, form trait–behavior clusters, and derive trait ratings from the implications of this concept as described by Srull and Wyer's (1989) model. The interpretation of several behaviors as exemplars of the same trait does not depend on the prototypicality status of the acts for a given trait concept.

Repeated exposure to the behaviors should result in a higher accessibility of the behavioral information for subsequent judgments. The identification of the target person who performed an act (Study 2), however, requires a parsimonious representation of information in memory. Interpretations of the act descriptions in terms of a single trait concept and the formation of trait–behavior clusters facilitates the initial learning of associations between targets and acts. Trait behavior clusters may be further consolidated by the additional learning trials. Thus, the training procedure likely results in a neglect of the multiple meanings of the behavioral information at encoding, which seems not to be compensated at later stages of the process.

Our experimental procedures do not allow us to decide which of the manipulations is necessary or sufficient to produce a change from behavior-to-trait processing to trait-to-trait processing. The absence of reliable effects for the number of acts in the descriptions and the delay between the presentation of act descriptions and trait ratings emphasizes the importance of manipulations affecting the encoding of behavioral information.

We have chosen to vary task demands in combination between the two experiments to simulate different modes of data collection in psychological research. Study 1 mimics on-line codings of behaviors where raters are confronted with a sequence of related behaviors performed by a single target and provide trait ratings immediately after the observation of behaviors (e.g., the observation of a target's behavior in a situation). The conditions of Study 2 are far more typical for *retrospective rating tasks* where subjects often observe naturally occurring events (e.g., family or group interactions) and provide trait ratings after they have acquired impressions of several target persons (e.g., Borkenau & Ostendorf, 1987; Shweder & D'Andrade, 1980).

A serious weakness of our trait-to-trait model as well as of Shweder's and D'Andrade's (1979, 1980; D'Andrade, 1965, 1974; Shweder, 1982) distortion hypothesis is that both do not specify which concept will be chosen as an anchor point for subsequent judgments. Obviously, subjects anchor their judgments at a concept that closely matches the meaning of the act descriptions. Thus, our selection of the concept for which the acts had been nominated by an independent sample of subjects resulted in a close correspondence between the measure of the trait-to-trait process and the trait ratings. This selection, not surprisingly, seems more appropriate if the prototypicality status of the selected act descriptions is high than if it is low. As many behaviors can be interpreted in terms of even conceptually dissimilar trait concepts (Gergen, Hepburn, & Comer-Fisher, 1986) the identification of an anchor may require the collection of additional data, independent of the trait ratings, when heterogeneous stimuli are used. Expectancies concerning a target's behavior as well as the accessibility of trait concepts during encoding (Bargh, 1984; Higgins & King, 1981; Srull & Wyer, 1989; Wyer & Srull, 1981, 1986) may have a substantial impact on the subsequent anchoring of trait ratings.

Turning to the discussion of the systematic distortion hy-

pothesis, our data lend support to the weak version of D'Andrade (1965, 1974) and Shweder's (1982; Shweder & D'Andrade, 1979, 1980) claim (Borkenau, 1990; Funder, 1987): Trait ratings have been shown to be guided by the conceptual similarity among trait concepts in the way predicted by Shweder and D'Andrade. Under conditions that simulate retrospective ratings the alternative explanation that trait ratings reflect the multiple prototypicality of acts for the concepts under study can be ruled out. Given the artificial nature of our rating task, we have to be cautious in generalizing our results beyond the presentation of written act descriptions to real life trait inferences. In both studies subjects had little information on the targets. The information that was presented was homogeneous in that all act descriptions presented for a target had been nominated for a single trait concept. In addition, we used a highly redundant set of interpersonal traits in our experiments that conform closely to a circumplex structure (Gutmann, 1988; Wiggins, 1979; Wiggins, Phillips, & Trapnell, 1989).

Study 2 shows that trait-to-trait processing occurs at least when behaviors and traits are sampled within a single behavioral domain. This implies that models of memory-based trait ratings should explicitly incorporate persons' knowledge about conceptual links among trait concepts and qualifies the prominent role assigned to a general evaluative concept of the person (e.g., in Srull & Wyer's, 1989, model). Reference to a general evaluative concept may be relevant for less redundant sets of trait categories, which have been sampled across heterogeneous behavioral domains (e.g., interpersonal behavior vs. cognitive abilities).

We conclude that our results demonstrate a high degree of flexibility in processing behavioral information. Depending on task demands, subjects use at least two strategies to derive trait ratings from behavioral information. Our future research efforts will be directed at the generalizability of our results to more natural behavioral information and heterogeneous behavioral domains.

References

Angleitner, A., Buss, D. M., & Demtröder, A. I. (1990). A cross-cultural comparison using the act frequency approach (AFA) in West Germany and the United States. *European Journal of Personality, 4,* 187–207.

Angleitner, A., & Demtröder, A. (1988). Acts and dispositions: A reconsideration of the act frequency approach. *European Journal of Personality, 2,* 121–141.

Bargh, J. A. (1984). Automatic and conscious processing of social information. In R. S. Wyer & T. K. Srull (Eds.), *Handbook of social psychology* (Vol. 3, pp. 1–43). Hillsdale, NJ: Erlbaum.

Berman, J. S., & Kenny, D. A. (1976). Correlational bias in observer ratings. *Journal of Personality and Social Psychology, 34,* 263–273.

Berman, J. S., Read, S. J., & Kenny, D. A. (1983). Processing inconsistent social information. *Journal of Personality and Social Psychology, 34,* 263–273.

Block, J. (1977). Correlational bias in observer ratings: Another perspective on the Berman and Kenny study. *Journal of Personality and Social Psychology, 35,* 873–880.

Block, J., Weiss, D. S., & Thorne, A. (1979). How relevant is a semantic similarity interpretation of personality ratings? *Journal of Personality and Social Psychology, 37,* 1055–1074.

Borkenau, P. (1986a). Systematic distortions in the recognition of trait information. In A. Angleitner, A. Furnham, & G. L. van Heck (Eds.), *Personality psychology in Europe: Current trends and controversies* (Vol. 2, pp. 177–191). Lisse, The Netherlands: Swets & Zeitlinger.

Borkenau, P. (1986b). Toward an understanding of trait interrelations: Acts as instances for several traits. *Journal of Personality and Social Psychology, 51,* 371–381.

Borkenau, P. (1990). Systematic distortion and systematic overlap in personality ratings. In G. van Heck, S. Hampson, J. Reykowski, & J. Zakrzewski (Eds.), *Personality psychology in Europe: Foundations, models, and inquiries* (Vol. 3, pp. 3–29). Amsterdam: Swets & Zeitlinger.

Borkenau, P., & Ostendorf, F. (1987). Retrospective estimates of act frequencies: How accurately do they reflect reality? *Journal of Personality and Social Psychology, 52,* 626–638.

Bourne, E. (1977). Can we describe an individual's personality? Agreement on stereotype versus individual attributes. *Journal of Personality and Social Psychology, 35,* 863–872.

Bruner, J. S., & Tagiuri, R. (1954). The perception of people. In G. Lindzey (Ed.), *Handbook of social psychology* (Vol. 2, pp. 634–654). Reading, MA: Addison-Wesley.

Buss, D. M. (1981). *The act frequency analysis of interpersonal dispositions.* Unpublished doctoral dissertation, University of California, Berkeley.

Buss, D. M. (1985, October). *The interpersonal categorization of acts.* Paper presented at the meeting of the Society for Experimental Social Psychology, Chicago, IL.

Buss, D. M., & Craik, K. H. (1980). The frequency concept of disposition: Dominance and prototypical dominant acts. *Journal of Personality, 48,* 379–392.

Buss, D. M., & Craik, K. H. (1983). The act frequency approach to personality. *Psychological Review, 90,* 105–126.

Buss, D. M., & Craik, K. H. (1984). Acts, dispositions, and personality. In B. A. Maher & W. B. Maher (Eds.), *Progress in experimental personality research: Normal personality processes* (Vol. 13, pp. 240–300). San Diego, CA: Academic Press.

Buss, D. M., & Craik, K. H. (1986). The act frequency approach and the construction of personality. In A. Angleitner, A. Furnham, & G. Van Heck (Eds.), *Personality psychology in Europe: Current trends and controversies* (Vol. 2, pp. 141–156). Lisse, The Netherlands: Swets & Zeitlinger.

Cantor, N., & Mischel, W. (1977). Traits as prototypes: Effects on recognition memory. *Journal of Personality and Social Psychology, 35,* 38–48.

Cantor, N., & Mischel, W. (1979a). Prototypes in person perception. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 12, pp. 3–52), San Diego, CA: Academic Press.

Cantor, N., & Mischel, W. (1979b). Prototypicality and personality: Effects on free recall and personality impressions. *Journal of Research in Personality, 13,* 187–205.

Chapman, L. J., & Chapman, J. P. (1967). Genesis of popular but erroneous psychodiagnostic observations. *Journal of Abnormal Psychology, 72,* 193–204.

Chapman, L. J., & Chapman, J. P. (1969). Illusory correlations as an obstacle to the use of valid psychodiagnostic signs. *Journal of Abnormal Psychology, 74,* 271–280.

Cohen, C. E. (1981a). Goals and schemata in person perception: Making sense from the stream of behavior. In N. Cantor & J. F. Kihlstrom (Eds.), *Personality, cognition, and social interaction* (pp. 45–68). Hillsdale, NJ: Erlbaum.

Cohen, C. E. (1981b). Person categories and social perception: Testing the boundaries of the processing effects of prior knowledge. *Journal of Personality and Social Psychology, 44,* 55–66.

Cooper, W. H. (1981). Ubiquitous halo. *Psychological Bulletin, 90,* 218–244.

D'Andrade, R. G. (1965). Trait psychology and componential analysis. *American Anthropologist, 67,* 215–228.

D'Andrade, R. G. (1974). Memory and the assessment of behavior. In H. M. Blalock (Ed.), *Measurement in the social sciences* (pp. 159–186). Chicago, IL: Aldine-Atherton.

DeSoto, C. B., Hamilton, M. M., & Taylor, R. B. (1985). Words, people, and implicit personality theory. *Social Cognition, 3,* 369–382.

Epstein, S. (1979). The stability of behavior: I. On predicting most of the people much of the time. *Journal of Personality and Social Psychology, 37,* 1097–1126.

Funder, D. C. (1987). Errors and mistakes: Evaluating the accuracy of social judgment. *Psychological Bulletin, 101,* 75–90.

Gergen, K. J., Hepburn, A., & Comer-Fisher, D. (1986). Hermeneutics of personality description. *Journal of Personality and Social Psychology, 50,* 1261–1270.

Gutmann, M. (1988). *Systematische Beziehungen interpersoneller Eigenschaften* [The structure of interpersonal traits]. Regensburg, Germany: Roderer.

Hastie, R., & Park, B. (1986). The relationship between memory and judgment depends on whether the judgment task is memory-based or on-line. *Psychological Review, 93,* 258–268.

Hastie, R., & Pennington, N. (1989). Notes on the distinction between memory-based versus on-line judgment. In J. N. Bassili (Ed.), *On-line cognition in person perception* (pp. 1–17). Hillsdale, NJ: Erlbaum.

Higgins, E. T., & King, G. (1981). Accessibility of social constructs: Information processing consequences of individual and contextual variability. In N. Cantor & J. F. Kihlstrom (Eds.), *Personality, cognition and social interaction* (pp. 60–121). Hillsdale, NJ: Erlbaum.

Jackson, D. N., Chan, D. W., & Stricker, L. J. (1979). Implicit personality theory: Is it illusory? *Journal of Personality, 47,* 1–10.

Lamiell, J. T., Foss, M. A., & Cavenee, P. (1980). On the relationship between conceptual schemes and behavior reports. *Journal of Personality, 48,* 54–73.

Mervis, C. B., & Rosch, E. (1981). Categorization of natural objects. *Annual Review of Psychology, 32,* 89–115.

Mischel, W. (1968). *Personality and assessment.* New York: Wiley.

Moskowitz, G. B., & Roman, R. J. (1992). Spontaneous trait inferences as self-generated primes: Implications for conscious social judgment. *Journal of Personality and Social Psychology, 62,* 728–738.

Newtson, D. A. (1976). Foundations of attribution: The perception of ongoing behavior. In J. H. Harvey, W. J. Ickes, & R. F. Kidd (Eds.), *New directions in attribution research* (Vol. 1, pp. 223–247). Hillsdale, NJ: Erlbaum.

Nisbett, R. E., & Ross, L. (1980). *Human inference: Strategies and shortcomings.* Englewood Cliffs, NJ: Prentice-Hall.

Park, B. (1989). Trait attributes as on-line organizers in person impressions. In J. N. Bassili (Ed.), *On-line cognition in person perception* (pp. 39–59). Hillsdale, NJ: Erlbaum.

Romer, D., & Revelle, W. (1984). Personality traits: Fact or fiction? A critique of the Shweder and D'Andrade systematic distortion hypothesis. *Journal of Personality and Social Psychology, 47,* 1028–1042.

Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General, 104,* 192–233.

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology, 7,* 573–605.

Rosenberg, S., & Sedlak, A. (1972). Structural representations of implicit personality theory. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 6, pp. 235–297). San Diego, CA: Academic Press.

Rothbart, M., Fulero, S., Jensen, C., Howard, J., & Birrell, P. (1978). From individual to group impressions: Availability heuristics in stereotype formation. *Journal of Experimental Social Psychology, 14,* 237–255.

Schneider, D. J. (1973). Implicit personality theory: A review. *Psychological Bulletin, 79,* 294–309.

Shweder, R. A. (1975). How relevant is an individual difference theory of personality. *Journal of Personality, 43,* 455–484.

Shweder, R. A. (1982). Fact and artifact in trait perception: The systematic distortion hypothesis. In B. A. Maher (Ed.), *Progress in experimental personality research* (Vol. 11, pp. 65–100). San Diego, CA: Academic Press.

Shweder, R. A., & D'Andrade, R. G. (1979). Accurate reflection or systematic distortion? A reply to Block, Weiss, and Thorne. *Journal of Personality and Social Psychology, 37,* 1075–1084.

Shweder, R. A., & D'Andrade, R. G. (1980). The systematic distortion hypothesis. In R. A. Shweder (Ed.), *Fallible judgment in behavioral research. New directions for methodology of social and behavior science* (Vol. 4, pp. 37–58). San Francisco, CA: Jossey-Bass.

Smith, E. R. (1989). Procedural efficiency and on-line social judgments. In J. N. Bassili (Ed.), *On-line cognition in person perception* (pp. 19–37). Hillsdale, NJ: Erlbaum.

Srull, T. K., & Wyer, R. S. (1989). Person memory and judgment. *Psychological Review, 96,* 58–83.

Weiss, D. S. (1979). The effects of systematic variations in information on judges' descriptions of personality. *Journal of Personality and Social Psychology, 37,* 2121–2136.

Weiss, D. S., & Mendelsohn, G. A. (1986). An empirical demonstration of the implausibility of the semantic similarity explanation of how trait ratings are made and what they mean. *Journal of Personality and Social Psychology, 50,* 595–601.

Wiggins, J. S. (1979). A psychological taxonomy of trait-descriptive terms: The interpersonal domain. *Journal of Personality and Social Psychology, 37,* 395–412.

Wiggins, J. S., Phillips, N., & Trapnell, P. (1989). Circular reasoning about interpersonal behavior: Evidence concerning some untested assumptions underlying diagnostic classification. *Journal of Personality and Social Psychology, 56,* 296–305.

Winter, L., & Uleman, J. S. (1984). When are social judgments made? Evidence for the spontaneousness of trait inferences. *Journal of Personality and Social Psychology, 47,* 237–252.

Wyer, R. S., & Srull, T. K. (1981). Category accessibility: Some theoretical and empirical issues concerning the processing of social stimulus information. In E. T. Higgins, C. P. Herman, & M. P. Zanna (Eds.), *Social cognition: The Ontario Symposium* (Vol. 1, pp. 161–197). Hillsdale, NJ: Erlbaum.

Wyer, R. S., & Srull, T. K. (1986). Human cognition in its social context. *Psychological Review, 93,* 322–359.

Received December 20, 1991
Revision received September 11, 1992
Accepted September 16, 1992 ■