Identifying the Situationally Variable Subject: Correspondence among Different Self-Report Formats

Robert G. Turner and Bob J. Gilliam
Seaver College, Pepperdine University

The present study compared the results obtained from three different procedures for obtaining self-reports of behavioral consistency versus inconsistency: (1) the traditional bipolar rating scale; (2) the Nisbett, Caputo, Legant, and Maracek (1973) procedure whereby subjects check the term, its antonym or the phrase "it depends on the situation"; and (3) subject-generated lists of self-descriptive traits. Results showed a moderate association between self-reports of situational variability and central responses on the scaled format; however, omission of a term in self-generated lists was not strongly associated with either central responses on the scale format or situational responses on the inventory formulated according to Nisbett et al.

Bem and Allen (1974) concluded that a trait measure will be predictively valid only for the subset of subjects that is behaviorally consistent across situations on that dimension. They suggested that those subjects who are inconsistent on the dimension would have to be excluded from predictive research, since their trait-relevant behavior would not be predictable. In subsequent consistency research, different procedures have been employed to differentiate the situationally consistent subject from the inconsistent subject. One of the most common procedures has been to ask subjects to check the al-

ternative on lists of bipolar trait terms that is the most accurate descriptor of their behavior: a trait term (e.g., *assertive*), its antonym (e.g., *unassertive*), or the phrase "it depends on the situation" (Nisbett, Caputo, Legant, & Maracek, 1973). Those checking the "it depends on the situation" alternative are then labeled as inconsistent or situationally variable on that trait dimension (McGee & Snyder, 1975; Turner, 1978).

Another procedure to differentiate the behaviorally consistent from the inconsistent is to ask subjects to generate their own list of self-relevant descriptors. The results of Turner and Gilliland (in press) showed that subjects who did not employ in their self-descriptions either *assertive* or its antonym, *unassertive*, behaved similarly to self-described unassertive subjects in an ambiguous situation, but similar to self-described assertive subjects in a structured situation in which they were told to be forceful. At least in this study, therefore, the omission of a term from a self-generated list of applicable self-descriptors identified an area highly susceptible to situational influence.

It is not known whether these two procedures identify the same subjects as inconsistent or situationally variable on a particular trait dimension. For example, is the omission of a term in a self-generated list equivalent to a situational response on the Nisbett et al. format? In addition,

it is not known how these two procedures compare with the traditional procedure of having subjects mark on a 1 to 7 scale where they see themselves on a bipolar dimension. Since the end points of the scale are anchored by the term and its antonym, it might be argued than those marking near the center of the scale (Points 3, 4, and 5) are reporting themselves to be inconsistent on the trait. If this is the case, the Nisbett et al. procedure does not obtain any additional information than can be obtained from traditional scaling procedures.

A comparison of self-report formats might have implications for the report of Sorrentino and Short (1977) on the inconsistent behavior of subjects identified as moderates on measures of achievement and affiliation motivation. If midpoint or moderate responses on personality scales are equivalent to reports of situational variability, then the behavior of these subjects should be more inconsistent than those marking end points on the scales (Bem & Allen, 1974).

Purpose

The purpose of the present study was to compare the results obtained from three different procedures for obtaining self-reports of behavioral consistency and inconsistency: (1) the traditional bipolar rating scales; (2) the Nisbett et al. (1973) procedure whereby subjects check the term, its antonym, or the phrase "it depends on the situation"; and (3) subject-generated lists of self-descriptive terms.

A secondary purpose was to compare the responses obtained on the Nisbett inventory to those obtained by Goldberg (1978) for single trait terms. In that study, subjects checked whether each trait term was an accurate description, an inaccurate description, or only occasionally accurate, that is, depending on the situation. Since in the present study subjects responded to bipolar traits (i.e., a term, its antonym, or a situational response), an opportunity to determine the effects of different formats on the proportion of situational responses was possible.

Method

During the first week of the trimester, students in an introductory social sciEnce class (N= 195) completed a Self-Description Form asking them to list the traits that best described themselves. After this form was collected, the students responded to a list of 30 pairs of bipolar traits anchoring 7-point rating scales and to another 30-item inventory in the form of Nisbett et al. (1973) in which students checked as self-descriptive a term, its antonym, or the statement "it depends on the situation." Of the 30 bipolar traits, 20 were on both inventories. In an effort to cover this duplication, the 10 filler traits were different for the 2 inventories. These 20 terms were from those categories of the Layman-Donald trait taxonomy (Goldberg, 1975; 1977) that subjects most commonly employed in their self-generated descriptions in the Turner and Gilliland (in press) study. The order of the rating scales and Nisbett et al. inventories was couterbalanced over the total sample of students.

The social desirability ratings of the bipolar traits were obtained from the ratings reported by Norman (1967) on a scale from 1 (low desirability) to 9 (high desirability), based on the responses of 100 subjects (50 males and 50 females).

The consistency of students' self-generated lists of relevant self-descriptors was evaluated by employment of the Layman-McDonald trait taxonomy (Goldberg, 1977). These researchers have developed a taxonomy of 45 semantically related clusters of approximately 1,700 trait terms. Consistent with Bem and Allen's (1974) statement that any trait term is relevant to only a limited subset of people, Turner and Gilliland (in press) have reported that only a few of these semantically related clusters are employed by more than 50% of subjects in generating lists of self-relevant descriptors. Therefore, in evaluating the consistency of self-generated descriptions with those on the scaled and Nisbett et al. formats, only five of the most commonly employed trait clusters as identified by Turner and Gilliland were included. Each of these clusters con-

tained one of the 20 bipolar traits on the scaled and Nisbett et al. inventories. These five traits were *optimistic-pessimistic, patient-impatient, self-assured-insecure, assertive-unassertive,* and *outgoing-reserved.*

Using the Layman-McDonald taxonomy to identify trait terms semantically similar to these five bipolar traits, two raters independently evaluated students' lists of self-generated descriptors. Each rater determined whether a student's list contained (1) a semantic equivalent of the trait term (e.g., *optimistic*), (2) its antonym (e.g., *pessimistic*), (3) both the term and its antonym (e.g., *optimistic* and *pessimistic*), or (4) no equivalent of the bipolar trait. Inter-rater agreement on these categorizations ranged from 70% for patient-impatient to 96% for self-assured-insecure ($\overline{X}$ = 86%). In comparisons of self-generated descriptors with responses to the scaled and Nisbett et al. inventories, only those responses

on which the raters concurred as to the appropriate categorization were included in these analyses.

Results

Table 1 shows the number of students checking either the trait term, its antonym, or the phrase "it depends on the situation" on the inventory developed via the Nisbett et al. (1973) format. The proportion of students checking the category "it depends on the situation" varied from .38 for *reflective-impulsive* to .11 for *devious-sincere*. The mean proportion of responses in this category was .27. The number of students checking this category was negatively correlated with the difference in the social desirability ratings of the bipolar terms ($r = -.49$, $df = 18$, $p < .02$). Thus, if the bipolar terms had approximately equal social desirability ratings (*reflec-*

Table 1
Proportion of Subjects Checking the First Trait Term, the Second Term
or Depends on the Situation on Each of the Twenty Trait Dimensions
(N = 195)

	Dimension	First Term	Second Term	Depends on Situation
1.	Reflective–Impulsive	.38	.24	.38
2.	Sarcastic–Sympathetic	.12	.52	.36
3.	Insecure–Self Assured	.14	.52	.34
4.	Patient–Impatient	.49	.18	.33
5.	Cautious–Bold	.46	.21	.33
6.	Flexible–Stubborn	.51	.17	.32
7.	Assertive–Unassertive	.54	.14	.32
8.	Outgoing–Reserved	.53	.15	.31
9.	Anxious–Carefree	.40	.29	.31
10.	Tactful–Brash	.38	.31	.31
11.	Selfish–Unselfish	.15	.55	.30
12.	Thrifty–Unthrifty	.43	.28	.29
13.	Opinionated–Openminded	.15	.62	.23
14.	Subjective–Logical	.15	.62	.23
15.	Energetic–Unenergetic	.74	.05	.21
16.	Ambitious–Nonchalant	.71	.09	.20
17.	Organized–Disorganized	.69	.11	.20
18.	Optimistic–Pessimistic	.69	.13	.18
19.	Conscientious–Lackadaisical	.80	.03	.17
20.	Devious–Sincere	.05	.84	.11

tive-impulsive differed by 1.86 points), more students checked the "situation" category. When the terms differed considerably (*devious-sincere* differed by 5.42 points), more students checked the socially desirable alternative and fewer checked the "situation" category.

Table 2 presents a cross-tabulation of students' responses to the 20 bipolar traits on the 1 to 7 rating scale and the Nisbett et al. (1973) format. When students checked the "situation" category on the Nisbett et al. format, approximately two-thirds (68%) of the time they checked

the central points on the 1 to 7 scale (Points 3, 4, 5) for the same bipolar terms. The respective percentages for these points were 25%, 23%, and 20%, which sum to 68%. Thus, the majority of students checking the "situation" category on the Nisbett et al. format described themselves on the same bipolar terms by marking central points on the 1 to 7 rating scale. However, results for the converse were not as direct, i.e., if students described their position on the rating scale by a central number (Points 3, 4, 5), they did not necessarily check the "situation" cate-

Table 2
Cross-Tabulation of Responses to the Rating Scale
Versus the Nisbett, et at. (1973) Format

| Rating Scale | Nisbett Format | | | |
	More Desirable Term	Less Desirable Term	Depends on Situation	Total
1	606 .89 .38	23 .03 .02	53 .08 .05	682 1.00 .18
2	563 .70 .35	62 .08 .05	175 .22 .17	800 1.00 .21
3	270 .44 .17	74 .12 .06	264 .43 .25	608 1.00 .16
4	77 .20 .05	73 .19 .06	239 .61 .23	389 1.00 .10
5	49 .10 .03	229 .47 .20	205 .42 .20	483 1.00 .13
6	27 .06 .02	359 .77 .31	78 .17 .07	464 1.00 .12
7	20 .05 .01	333 .87 .29	29 .08 .03	382 1.00 .10
Total	1612 .42 1.00	1153 .30 1.00	1043 .27 1.00	3808 1.00 1.00

Note. Table entries are number of responses, row proportions and column proportions respectively.

gory on the Nisbett et al. format. Of the 608 responses on Point 3 of the 1 to 7 scale, 264 (or 43%) checked the "situation" category on the Nisbett et al. format. Of those checking Points 4 or 5 on the scale format, 61% and 42%, respectively, checked the "situation" category. For the three middle points on the scale format, 48% checked the "situation" category on the Nisbett et al. format.

The cross-tabulation of responses on the 1 to 7 rating scale with responses on the student-gen-erated trait lists is presented in Table 3. For the five pairs of bipolar traits included in this analysis, the more socially desirable of each pair was associated with low scores on the rating scale, while the less desirable was associated with the larger scale number. As discussed earlier, the self-generated lists were evaluated as to whether they contained (1) the trait term or a semantic equivalent of the term, (2) its antonym, (3) both the term and its antonym, or (4) neither the trait nor its antonym.

Table 3
Cross–Tabulation of Responses on Subject Generated Trait Lists with Responses at Each Point on the Rating Scale for Five Traits

Rating Scale	More Desirable Term	Less Desirable Term	Both Term	Neither Term Listed	Total
1	78 .42 .41	7 .04 .08	3 .02 .17	99 .53 .19	187 1.00 .23
2	56 .25 .29	13 .06 .14	4 .02 .22	149 .67 .28	222 1.00 .27
3	24 .16 .13	13 .09 .14	5 .03 .28	106 .72 .20	148 1.00 .18
4	10 .15 .05	11 .16 .12	1 .01 .06	46 .68 .09	68 1.00 .08
5	11 .12 .06	14 .15 .16	1 .01 .06	68 .72 .13	94 1.00 .11
6	9 .14 .05	16 .26 .18	1 .02 .06	36 .58 .07	62 1.00 .08
7	2 .05 .01	16 .36 .18	3 .07 .17	22 .51 .04	43 1.00 .05
Total	190 .23 1.00	90 .11 1.00	18 .02 1.00	526 .64 1.00	824 1.00

Note.   Table entries are number of responses, row proportions and column proportions respectively.

In general, the data presented in Table 3 show a strong influence of the relative social desirability levels of the bipolar traits in the rating scale. When students generated the trait term in their descriptions, they tended to respond consistently on the rating scale: 41% at Point 1, 30% at Point 2, or 71% total. Only 6% of responses were on Points 6 and 7 at the opposite, or inconsistent, end of the scale. However, when students employed the less socially desirable term in their self-generated lists, only 18% of responses were on each of Points 6 and 7 on the rating scale (total = 36%), while 22% of responses (8% + 14%) were at Points 1 and 2 at the opposite (inconsistent) end of the scale. In addition, if students employed both terms or did not employ either the trait term or its opposite in their descriptions, their responses to these terms in the rating scale tended to cluster at the points indicating the applicability of the socially desirable trait. Thus, the omission of a term in students' self-generated lists of personal descriptors did not tend to translate into responses at the central points on the rating scale.

As shown in the bottom row of Table 3, the majority of responses to the bipolar traits in the rating scale were paired with the "not listed" category in the self-generated lists. This row of percentages shows a tendency for the highest percentages to be associated with responses to central points (Points 3, 4, 5) on the scaled format (72%, 68%, 72%, respectively) and the lowest percentages (53% and 51%) to be associated with the extreme points (1 and 7, respectively).

Thus, the results presented in Table 3 might be summarized in the following way. Given that a person has not listed a trait term as a relevant self-descriptor, there was no tendency to respond to that trait with a central rating. However, given a central rating, there was a tendency for more students to omit the trait in their self-generated lists than if their response on the rating scale had been at the extreme points (1 or 7).

In Table 4 are presented the results of the cross-tabulation of responses to the Nisbett et al. format and the self-generated trait lists. If a person generated a term as a relevant descriptor, 73% of the time he/she attributed the trait to himself/herself on the Nisbett et al. format. If the person listed the less desirable antonym as a descriptor, 40% of the time he/she attributed this term to himself/herself on the Nisbett et al. format. These consistency percentages for the trait and its opposite were similar to those reported earlier for the cross-tabulation of the self-generated lists with the rating scale.

The failure to list a trait as a relevant self-descriptor was not equivalent to a "situation" response on the Nisbett et al. format. If a trait was not self-generated, 31% of the corresponding responses on the Nisbett et al. format were "it depends on the situation." This percentage is only slightly higher than the 27% average of situational responses associated with the other three student-generated list categories (21%, 40%, 22% for the trait, its antonym, and both, respectively). Given that the person checked the situational category on the Nisbett et al. format, 67% of the time the trait was not self-generated. However, if the trait or its opposite were checked, 57% and 75% of the time, respectively ($\bar{X}$ = 66%), the terms were not self-generated. Therefore, there was no indication of an association between omission of a trait in a self-generated list of descriptors and responses of situational variability on that trait on the Nisbett et al. format.

Discussion

One of the main purposes of the present study was to determine the extent to which self-reported situational variability on a bipolar trait as identified by the Nisbett et al. (1973) procedure was equivalent to endorsement of cental points on a 1 to 7 rating scale anchored by the bipolar terms. Given a "situation" response to a bipolar trait on the Nisbett et al. format, a majority of the time (68%) the corresponding responses to the same bipolar trait on the rating scale was a central point (Point 3, 4, or 5). However, given a central position on the rating scale,

Table 4
Cross-Tabulation of Responses on Subject Generated Trait Lists with
Responses on the Nisbett et al. Format

| Subject Generated Lists | Nisbett Format | | | |
	More Desirable Term	Less Desirable Term	Depends on Situation	Total
More Desirable Term	141 .73 .36	10 .05 .05	41 .21 .17	192 1.00 .23
Less Desirable Term	18 .21 .05	35 .40 .19	35 .40 .14	88 1.00 .02
Both Terms	12 .67 .03	2 .11 .01	4 .22 .02	14 1.00 .02
Neither Term	225 .42 .57	141 .27 .75	164 .31 .67	530 1.00 .64
Total	396 .48 1.00	188 .23 1.00	244 .29 1.00	828 1.00 1.00

Note.   Table entries are number of responses, row proportions and column proportions respectively.

only for the most central point (Point 4) was a majority (61%) of the corresponding responses on the Nisbett et al. format in the "situation" category.

Although the correspondence between responses to the "situation" category on the Nisbett et al. format and responses at the central points on the scale format was not as high as might be desired, the extent of the correspondence (48% average for Points 3, 4, and 5) is perhaps sufficient to apply to the report of Sorrentino and Short (1977) on the inconsistent behavior of subjects identified as moderates on measures of achievement and affiliation motivation. Based upon data from four studies and an extensive review of the motivation literature, these authors concluded that there was "a pervasive inconsistency in the behavior of those who obtain moderate scores on the various motive measures" (p. 478). In discussing possible explanations for this inconsistency, one of the alternatives Sorrentino and Short offered was that these moderate-scoring subjects may be equivalent to the situationally variable subjects discussed by Bem and Allen (1974). In support of this explanation, the present results would suggest that approximately one-half of these moderate-scoring subjects would have reported themselves as situationally variable on the motivation dimensions employed in the reviewed studies.

In general, there was no strong trend for the omission of a trait as a relevant self-descriptor on the self-generated trait lists to be paired with either central responses to that trait on the rating scale or with reports of situational variability on the trait on the Nisbett et al. format. In an investigation of assertive behavior, Turner and Gilliland (in press) reported that subjects not listing semantic equivalents of assertiveness or unassertiveness in their self-descriptions behaved similarly to the unassertive subjects in an unstructured group problem-solving task, while they behaved similarly to assertive subjects in a similar situation in which they were told to be dominant. A synthesis with the present results would suggest that although omission of a trait term is not consistent with a response of situational variability on the trait, it does seem to be associated with behavioral variability, at least for assertiveness.

Relevant to the results shown in Table 1 for the Nisbett et al. (1973) inventory, Goldberg (1978) comprehensively investigated responses to single trait terms where the response alternatives were (1) the word is *not* a good or accurate description; (2) the word *is* a good or accurate description; or (3) the word is only occasionally descriptive—in other words, if it depends on the situation. In describing themselves, Goldberg's subjects employed the situational response approximately .46 of the time. In the present study where the trait term, its antonym, or "it depends on the situation" were the alternatives, only .27 of the responses were in the situational category.

Although it is possible that this difference in proportions resulted from sample differences in respondents and/or the trait terms employed, it seems more likely that the presence of antonyms of the terms used in the present study led to a reduction in the proportion of situational responses. For example, in confronting the item *anxious*, a subject might check the "accurate" category or "not accurate" category if the term is either highly self-descriptive or if it is almost never self-descriptive. Assuming that the subject is aware of variabilities in the extent to which

either extreme is true, a situational response is encouraged. However, the addition of the antonym *anxious-carefree* encourages the subject to select the alternative that *tends* to be more accurate, leaving the situational response for those instances that are too close to call. A smaller proportion of situational responses should, therefore, be expected for bipolar items compared to single term items.

This theoretical integration of the present results with those of Goldberg (1978) assumes the absence of response styles for simplicity of presentation. Both studies showed utilization of the situational response alternative to be related to the social desirability of the trait terms. In Goldberg (1978) the situational response was used more often when the term had a social desirability rating near the middle of the scale. In the present study, the situational response tended to occur when the term and its antonym had similar social desirability ratings. Both results showed, therefore, that utilization of the situational response tended to occur when subjects were not missing a chance to assign a highly desirable characteristic to themselves or to deny the applicability of a highly undesirable one. Finally, it should be noted that only self-descriptions were employed in the present study. The description of other people on the different inventories may or may not result in cross-tabulations similar to those obtained here.

References

Bem, D. J., & Allen, A. On predicting some of the people some of the time: The search for cross-situational consistencies in behavior. *Psychological Review*, 1974, *81*, 506–520.

Goldberg, L. R. Toward a taxonomy of personality descriptive terms: A description of the O.R.I. taxonomy project. *Oregon Research Institute Technical Report*, 1975, *15*, No. 2.

Goldberg, L. R. *Language and personality: Developing a taxonomy of trait-descriptive terms.* Invited address presented at the annual meeting of the American Psychological Association, San Francisco, CA, 1977.

Goldberg, L. R. Differential attribution of trait-descriptive terms to oneself as compared to well-liked, neutral, and disliked others: A psychometric analysis. *Journal of Personality and Social Psychology,* 1978, *36,* 1012–1028.

McGee, M. G., & Snyder, M. Attribution and behavior: Two field studies. *Journal of Personality and Social Psychology,* 1975, *32,* 185–190.

Nisbett, R. E., Caputo, C., Legant, P., & Maracek, J. Behavior as seen by the actor and as seen by the observer. *Journal of Personality and Social Psychology,* 1973, *27,* 154–164.

Norman, W. T. *2800 personality trait descriptors: Normative operating characteristics for a university population.* Unpublished manuscript, University of Michigan, Department of Psychology, April 1967.

Sorrentino, R. M., & Short, J. C. The case of the mysterious moderates: Why motives sometimes fail to predict behavior. *Journal of Personality and Social Psychology,* 1977, *35,* 478–484.

Turner, R. G. Consistency, self-consciousness, and the predictive validity of typical and maximal personality measures. *Journal of Research in Personality,* 1978, *12,* 117–132.

Turner, R. G., & Gilliland, L. The comparative relevance and predictive validity of subject-generated trait descriptors. *Journal of Personality,* in press.

Acknowledgments

Author's Address

Send requests for reprints or further information to Robert G. Turner, 660 Parrington Oval, Room 114, University of Oklahoma, Norman, OK 73019.