

**PERSONAL AND SITUATIONAL
DETERMINANTS OF PERSONALITY
RESPONSES: A PARTIAL REANALYSIS AND
REINTERPRETATION OF THE SCHMIT ET
AL. (1995) DATA**

Chet Robie

University of Houston

Marise Ph. Born

Vrije Universiteit Amsterdam

Mark J. Schmit

Personnel Decisions International

ABSTRACT: Using generalizability theory, a partial reanalysis and reinterpretation of the Schmit et al. (1995) data was conducted with undergraduate students who were asked to respond to both noncontextualized and work-specific versions of the NEO Five-Factor Inventory (Costa & McCrae, 1989) either honestly ($N = 100$) or as if they were applying for a job ($N = 100$). Results indicated that the person variance was *larger* in the applicant compared to the general instructions condition, that the person \times situation interaction accounted for significant variance across all scales included in the study (Neuroticism, Extraversion, Agreeableness, and Conscientiousness), that these person \times situation effects were attenuated in the applicant instructions condition as compared to the general instructions condition, and that responses to the Neuroticism scale were most affected by this interaction. Implications for the use of personality constructs in personnel selection include a shift from concerns of “faking” to a con-

We thank Beth Manke for her thoughtful comments.

A version of this paper was presented as part of a symposium conducted at the 9th European Conference on Personality, University of Surrey, Guilford, UK.

Address correspondence to Chet Robie, Niagara University, College of Business Administration, Niagara University, New York 14109; e-mail: chetrobie@gateway.net.

cern for proper contextualization of personality measures with the goal of maximizing prediction of work-related behavior.

KEY WORDS: Personality; personnel selection; interactionism; faking.

The consistency of individual behavior across situations has been of long-standing interest in personality psychology. Recent interactional research in the personality literature suggests that dispositions are conditional; that is, although there is a connection between personality and behavior, the power to predict behavior may be limited to a fairly specific range of situations (Mischel & Shoda, 1995; Murtha, Kanfer, & Ackerman, 1996; Roberts & Donahue, 1994; Wright & Mischel, 1987).

Recent research in organizational psychology has drawn from this literature in looking at the effects of situational influences on individual differences constructs that are used in personnel selection. A recent example of the use of an interactional approach to the use of personality constructs in personnel selection is the study by Schmit, Ryan, Stierwalt, and Powell (1995). Schmit et al. conducted two laboratory studies that investigated the effects of a situational manipulation (i.e., frame-of-reference) on personality scale scores and criterion-related validity. In Study 1, both within- and between-groups designs were employed to assess the effects of testing situation (general instructions versus applicant instructions) and item type (work specific vs. noncontextual) on personality scale scores (Neuroticism, Extraversion, Agreeableness, and Conscientiousness). Results indicated that a work-related testing context and work-related items led to more positive responses (i.e., lower mean scale scores for Neuroticism and higher mean scale scores for Extraversion, Agreeableness, and Conscientiousness). A second study was designed to examine the implications of these score differences on the criterion-related validity of one of the personality factors. This second study found differences in the predictive validity of several facets of Conscientiousness, depending on frame of reference and instructions. Specifically, context-specific items given in an applicant instructions condition were found to have the consistently highest predictive validity. The implications for the use of personality constructs in personnel selection seemed clear—the use of work specific items in personality scales used for personnel selection will increase the scale scores yet will also increase criterion-related validity, thereby presumably increasing the utility of personality constructs in personnel selection.

Schmit et al.'s reasoning for conducting the first study was that the mean differences in personality scale scores across the different instruction and item types, presumably caused by self-presentation processes (i.e., the process of presenting an image of how one hopes to be regarded

by others), could affect the criterion-related validity of those scales. In fact, at one point in the article, Schmit et al. note about their results, "For example, mean differences across item type conditions were not found for the Extraversion scale in both Substudies 1 and 2. This finding leaves open the possibility that validity differences comparable to those found for the Conscientiousness scale in Study 2 may not be found for a measure of extraversion" (p. 618). However, a mean difference (i.e., mean shift) may indicate a constant bias across individuals which would not affect validity. A mean shift across conditions could indicate that all respondents change similarly from one context to another (this is not likely to be true) but could also indicate both intercondition and interindividual differences in change.

Self-presentation probably cannot be modeled appropriately simply by examining mean differences across conditions that differ in environmental press; specifically, research suggests that individual differences in self-presentation styles exist (e.g., Olson & Johnson, 1991) and Hogan (1991) suggests that individuals may differ in their ability to self-present. Research further suggests that personality responses in situations that differ in environmental press may also be affected by an individual's level of self-monitoring, or the degree to which one is cross-situationally consistent (Snyder, 1983). That is, the theory that individuals have a situation specific personality (e.g., Murtha et al., 1996) probably has merit, but individuals will likely differ in the degree to which their personality is cross-situationally consistent or inconsistent. Thus, although a given environmental press may cause a mean shift in personality responses, it is also possible that responses for some individuals will be affected more than those of others (i.e., a differential shift). If self-presentation is a phenomenon that entails interindividual differences, it seems likely that Schmit et al. (1995) missed an opportunity to study the specific effects of frame-of-reference on personality responses.

A differential shift (i.e., person \times situation interaction) may have a more direct impact on validity in comparison to a mean shift (i.e., situation main effect) because a differential shift denotes a shift in rank order, which per statistical definition, will affect validity. Thus, the study of person \times situation effects in personality responses is highly consequential for personnel selection practice because, should the rank order of respondents shift across situations (i.e., frame-of-reference in the current study), criterion-related validity will change and ultimately the hiring decisions based on the measure will differ depending on the situational context used in the items of the measure. It should also be noted that a change in person variance denotes a change in the amount of construct variance which, also per statistical definition, will affect validity.

The current study is firstly an attempt to further explore the possible effects of frame-of-reference on the validity of four of the Big-Five

personality factors (Neuroticism, Extraversion, Agreeableness, and Conscientiousness) through a partial reanalysis and reinterpretation of the Schmit et al. (1995) data using the person \times situation experimental design (Endler & Magnusson, 1976) and generalizability theory (Shavelson & Webb, 1991) as major tools.¹ Specifically, we are interested in ascertaining how much of the variance in personality responses is attributable to the person, how much is due to the situation (i.e., item type or frame-of-reference), how much is due to the person \times situation interaction, and how much is due to measurement error. Also, we are interested in what effects a clear environmental press such as instruction type (i.e., general vs. applicant) has on these variance components. Because the effect of frame-of-reference on the criterion-related validity was only directly studied in relation to several facets of the Conscientiousness scale in the Schmit et al. study, our reanalysis and reinterpretation is, secondly, an attempt to model the possible effects of frame-of-reference on the criterion-related validity of the three other factors (Neuroticism, Extraversion, and Agreeableness).

Irrespective of instructional condition, we have several expectations for the results of this reanalysis. First, we expect significant person variance in the responses to the measures because years of personality research has convincingly provided evidence that persons do reliably differ along certain dispositional dimensions cross-situationally (Epstein & O'Brien, 1985). Second, we expect significant situation variance for the personality scales because a work specific context provides a more clear environmental press than a general context; moreover, Schmit et al.'s findings of mean differences in scale scores between item types for the Neuroticism, Agreeableness, and Conscientiousness scales should hold here (even though our analytic techniques were slightly different from that of Schmit et al.—see the Statistical Analyses section below). Third, we expect significant person \times situation variance because research in the interactional tradition has provided convincing evidence that interactions between persons and situations do exist (Endler & Edwards, 1986). However, the practical question we are attempting to answer goes beyond the issue of “do person \times situation interactions exist?” but instead focuses on how much variance is attributable to these interactions and which personality factors seem to be affected the most.

We believe that the person \times situations interactions will probably be larger for personality constructs that are more strongly related to measures of social desirability. Measures of social desirability appear to be more a function of individual differences and less a function of the biasing effects of the items or scales (McCrae & Costa, 1983; Ones, Viswes-

¹As noted in Schmit et al. (1995, p. 609), the Openness to Experience scale was not included in the study because work appendages did not make sense on many of the items.

varan, & Reiss, 1996). That is, it appears that social desirability measures more likely tap something akin to self-monitoring or within-situation trait variance. Thus, as discussed earlier, this type of measure has the potential to lead to person \times situation interactions through a differential shift in responses. Accordingly, the size of the relations between scores on social desirability measures and personality factors similar to the scales used in the present study can give us some direction in making predictions concerning the size of the interactions that we might expect for each personality scale in our study. A large-scale meta-analysis by Ones et al. (1996) found that the relations between social desirability and the Big-Five personality factors were highest for Emotional Stability (i.e., reverse-coded Neuroticism) ($\rho = .37$), followed by Conscientiousness ($\rho = .20$), Agreeableness ($\rho = .14$), Extraversion ($\rho = .06$), and Openness to Experience ($\rho = .00$), respectively. Our predictions for the sizes of the person \times situation effects in the present study therefore follow this order.

We also have several expectations for possible differences in variance components for the general versus the applicants instructions conditions. One current perspective on the use of personality variables in personnel selection endorses the belief that the fidelity of trait measurement is somehow eroded when personality measures are used in selection contexts (cf. Douglas, McDaniel, & Snell, 1996; Frei, Griffith, Snell, McDaniel, & Douglas, 1997; Zickar, 1997). A study by Christiansen (1998), which seems to be prototypical of this perspective, suggests that trait variance is reduced in an applicant in comparison to a general instructions condition. In his study, 400 psychology undergraduates were asked to honestly fill out a personality questionnaire measuring Extraversion and Conscientiousness. Then, they were asked to fill out a second questionnaire with alternate measures of the same traits either honestly ($N = 200$) or as if they were applying for a sales job they wanted very much ($N = 200$). Christiansen (1998) found that the correlations between pre- and post-manipulation trait scores for Extraversion and Conscientiousness were significantly lower in the applicant instructions condition ($r = .42$ and $.43$, respectively) in comparison to the general instructions condition ($r = .74$ and $.65$, respectively) which he suggested was due to response distortion occurring in the applicant instructions condition and thus was evidence of a deterioration of trait variance in the applicant instructions condition.

Based on the results of the Christiansen (1998) study, one might expect the person variance to be reduced in the applicant instructions condition in our study. In addition one might expect that the situation and person \times situation effects would be diminished when respondents are asked to respond as if they were applying for a job. However, an alternative explanation for these types of effects is based on the theory that individuals have situational-dispositional traits (Murtha et al.,

1996). Under this theory, one would expect person variance to be increased in the applicant instructions conditions in our study because the general instructions induce a situational reference that is too broad. Further, when responding to a personality inventory as if one were applying for a job, individuals may already have an implicit work frame of reference which could act to negate any frame-of-reference person \times situation effect. Personality seems to evidence greater consistency and coherence when measured in similar situations (cf. Murtha et al., 1996); thus, the hypothesized decrease in the person \times situation interaction may not be due to an increase in error associated with "faking" as Christiansen (1998) suggested but may instead be due to a decrease in non-relevant situational noise.

Based on the research outlined above, the following predictions were made:

Hypothesis 1. Individuals will differ from one another with respect to their personality responses (a person main effect).

Hypothesis 2. Item type (i.e., situation) will affect personality responses such that the item type with the most clear environmental press (i.e., work specific) will evidence the most positive scale scores (a situation main effect).

Hypothesis 3. Individuals will differ in the extent to which their personality responses change across item types (a person \times situation interaction). This interaction will be largest for Neuroticism, followed by Conscientiousness, Agreeableness, and Extraversion, respectively.

Hypothesis 4. Instruction type will result in a change in the person, situation, and person \times situation effects in the applicant instructions condition in comparison to the general instructions condition. The amount and direction of change was not specified as tests of this hypothesis was considered an exploratory comparison of theories.

METHOD

Participants

Data from the second substudy of the first study in Schmit et al. was used. Students of an introductory psychology course participated for course credit. Data were collected from a total of 200 participants, with

100 participants randomly assigned to each of the two between-groups conditions. Across groups, participants did not differ significantly in age, class rank, or the number of jobs previously held.

Measures

The NEO Five-Factor Inventory (NEO-FFI; Costa & McCrae, 1989) was the personality measure used in the Schmit et al. study. The NEO-FFI is a shortened form of the NEO PI-R (Costa & McCrae, 1992), a measure based on the five factor model of personality. The five factors of personality that the NEO-FFI measures are Neuroticism, Extraversion, Openness to Experience, Agreeableness, and Conscientiousness. The NEO-FFI consists of the 12 items having the highest positive or negative loading on each of the corresponding factors on the NEO PI-R. Responses to each item are recorded on a 5-point scale ranging from 0 (*strongly disagree*) to 4 (*strongly agree*). Previous estimates of internal consistency (coefficient alphas) for each of the 12-item scales were .86 (Neuroticism), .77 (Extraversion), .73 (Openness to Experience), .68 (Agreeableness), and .81 (Conscientiousness) (Costa & McCrae, 1992). Costa and McCrae (1992) also reported that correlations between NEO-FFI scales and NEO PI-R factors ranged from .75 to .89.

The second personality inventory used was an altered form of the NEO-FFI, on which a reference to work was appended to each statement, usually at the beginning or the end of the statement. For example, the item “I try to be courteous to everyone I meet” was modified to read “I try to be courteous to everyone I meet at work.” Another example item was “I work hard to accomplish my work-related goals” instead of “I work hard to accomplish my goals.” As noted earlier, the Openness to Experience scale was not used in the Schmit et al. study (and thus will not be used in the current study) because work appendages did not make sense on many of the items.

Procedure

Participants were randomly assigned to one of two groups (general instructions group or applicant instructions group). Participants in both groups completed both forms (counterbalanced within groups) of the personality test. Written instructions were attached to each personality inventory. The general instructions group was instructed to answer the questions as directed on the original version of the test. The applicant instructions group was instructed to complete the inventories as if they were applying for a customer service representative in a department store, a job they really wanted.

Statistical Analyses

Consistent with the methodology reported in a recent generalizability study by Lakey, McCabe, Fiscaro, and Drew (1996), in our first set of analyses, we analyzed the data as a fully crossed mixed analysis of variance (ANOVA) with random factors.² Persons (i.e., participants) was the between-subjects variable, and situations (i.e., item type) and items were the within-subjects variables. Each participant served as a level of the person variable, each item type served as a level of the situation variable, and each item served as a level of the item variable. Thus, a 100 (persons) \times 2 (situations) \times 12 (items) ANOVA was employed. We conducted this analysis for each scale and each instruction type. Because this design produces only one observation per cell, the appropriate error term is the highest order interaction (Lindman, 1974; Shavelson & Webb, 1991). Our interest was not in item variance or variance attributable to the interactions of items with other factors. However, as noted by Lakey et al. (1996), unlike some prior applications of generalizability theory (cf. the eleven person \times situation studies reviewed in Bowers, 1973), residual variance using our design is not reflected in the person \times situation interaction. This is because treating items as an independent variable enables the person \times situation \times item interaction to capture random error. When such a three-way interaction effect is not present, this term provides an accurate index of random error. When such an interaction exists, the term is an overestimate of error, leading to more conservative significance tests (Lindman, 1974). Estimates of variance components were made following formulas provided in Shavelson and Webb (1991). Calculation of quasi F ratios to test the statistical significance of those variance components were made following formulas provided in Neter, Kutner, Nachtsheim, and Wasserman (1996).³ Because 48 tests of statistical significance were computed, to attain a familywise error rate of .05, we used an alpha of .001 (.05 / 48 \approx .001) to infer statistical significance.

In our second set of analyses, we attempted to model the variance components as they would occur for a typical selection scenario. Specifically, the variance components provided in a generalizability analysis reflect the variance attributable to each factor if only one level of that factor were randomly sampled. However, in a typical selection scenario, decisions made upon a given individual are not made at the item level

²The situation variable was considered random based on a pure traitest approach which would advocate interchangeability of the situations (i.e., regardless of the situation, individual differences should remain the same).

³Note that this analytic strategy differs markedly from that followed by Schmit et al. (1995). In that study, an instruction type \times item type mixed analysis of variance with fixed factors was employed.

but at the level of the scale score; moreover, individuals are not usually asked to complete the selection instrument using multiple frames of reference. Thus, we provided estimates of the variance components for a situation in which decisions are made about an individual based upon the mean scale score across 1 situation and 12 items. The formulas to do such analyses are found in Shavelson and Webb (1991). Because we were not manipulating the levels of the person and situation factors, the computations were straightforward; we simply divided the item main effect variance component and any factor that interacted with the item factor by 12 (the number of items in each scale). This procedure effectively acts to “wash out” much of the error due to item variance and allow us to examine the various sources of variance that would determine personality responses under a typical condition of measurement. As in our first set of analyses, we conducted this analysis for each scale and each instruction type.

RESULTS

The estimated variance components, percentages of variance, and quasi F ratios for personality responses by scale and instruction type are presented in Table 1. The person (i.e., true variance) factor and the person \times situation interaction accounted for significant variance in responses across all scales and both instruction types. However, the situation factor did not account for significant variance in responses for either instruction type for Extraversion or for the applicant instructions condition in the Neuroticism and Agreeableness factors.⁴ The item factor and person \times item interaction accounted for significant variance in all scales and conditions (with the exception of a nonsignificant item factor in the general instructions condition for the Conscientiousness scale). Moreover, the situation \times item interaction was not significant at all for either condition of the Extraversion scale and also for the general instructions condition for the Neuroticism and Agreeableness scales.

The mean scores estimated variance components and percentages of variance for personality responses by scale and instruction type for a typical selection scenario are presented in Table 2. Variance due to the item factor and its interactions with other factors (with the exception of the three-way interaction plus error term) was fairly consistently low (the highest was the Agreeableness factor), averaging approximately 9 percent of the total variance in personality responses across the scales

⁴Note that, when significant situation variance was evidenced, this was associated with more positive responses for the work specific item type (see Table 1, p. 610, Schmit et al., 1995).

Table 1
Estimated Random Effects Variance Components, Percentages of Variance,
and Quasi F Ratios for Personality Responses by Scale and Instruction Type

Scale	Variable	<i>General Instructions</i>			<i>Applicant Instructions</i>		
		$\hat{\sigma}^2(\alpha)$	% var	F	$\hat{\sigma}^2(\alpha)$	% var	F
Neuroticism	person	.21	15	3.09*	.28	24	5.12*
	situation	.13	9	56.11*	.02	2	10.98
	item	.16	11	20.15*	.09	8	9.79*
	p × s	.13	9	3.57*	.06	5	2.48*
	p × i	.18	13	1.61*	.18	16	1.68*
	s × i	.01	1	2.20	.01	1	3.18*
	p × s × i, error	.59	42		.51	44	
Extraversion	person	.21	20	4.11*	.16	18	4.15*
	situation	.00	0	0.03	.00	0	2.44
	item	.11	10	14.21*	.09	10	13.81*
	p × s	.06	6	2.67*	.03	3	2.06*
	p × i	.26	24	2.24*	.27	30	2.53*
	s × i	.01	1	2.57	.01	1	2.49
	p × s × i, error	.42	39		.35	38	
Agreeableness	person	.13	12	3.02*	.20	19	5.42*
	situation	.02	2	22.40*	.02	2	14.31
	item	.24	22	35.73*	.22	20	23.54*
	p × s	.04	4	2.57*	.02	2	1.58*
	p × i	.32	29	2.88*	.26	24	2.47*
	s × i	.01	1	2.21	.01	1	4.13*
	p × s × i, error	.34	31		.35	32	
Conscientiousness	person	.15	16	3.48*	.23	31	7.37*
	situation	.09	10	18.94*	.03	4	20.24*
	item	.07	8	3.41	.07	9	9.99*
	p × s	.07	8	3.50*	.03	4	2.41*
	p × i	.15	16	1.93*	.09	12	1.62*
	s × i	.05	5	15.55*	.01	1	4.39*
	p × s × i, error	.33	36		.29	39	

Notes. $N = 100$ for each instruction-type. * $p < .001$. $\hat{\sigma}^2(\alpha)$ = estimated random effects variance component for α . % var = percentage of variance accounted for (may not total 100% due to rounding error).

and instruction types. Variance due to the situation was more variable across the scales and instruction types, averaging approximately 10 percent; however, the situation did not account for any variance in the Extraversion scale and accounted for approximately one-quarter of the variance in the general instructions condition for the Neuroticism and Conscientiousness scales. The person × situation interaction more consistently accounted for a higher percentage of variance than any other term with the exception of the person variance—an average of 15 per-

Table 2
Mean Score Random Effects Estimated Variance Components and Percentages
of Variance for Personality Responses by Scale and Instruction Type
for a Typical Selection Scenario

Scale	Variable	<i>General Instructions</i>		<i>Applicant Instructions</i>	
		$\delta^2(\bar{\alpha})$	% var	$\delta^2(\bar{\alpha})$	% var
Neuroticism	person	.21	39	.28	67
	situation	.13	24	.02	5
	item	.01	2	.01	2
	p × s	.13	24	.06	14
	p × i	.01	2	.01	2
	s × i	.00	0	.00	0
	p × s × i, error	.05	9	.04	10
Extraversion	person	.21	64	.16	64
	situation	.00	0	.00	0
	item	.01	3	.01	4
	p × s	.06	18	.03	12
	p × i	.02	6	.02	8
	s × i	.00	0	.00	0
	p × s × i, error	.03	9	.03	12
Agreeableness	person	.13	48	.20	65
	situation	.02	7	.02	6
	item	.02	7	.02	6
	p × s	.04	15	.02	6
	p × i	.03	11	.02	6
	s × i	.00	0	.00	0
	p × s × i, error	.03	11	.03	10
Conscientiousness	person	.15	42	.23	70
	situation	.09	25	.03	9
	item	.01	3	.01	3
	p × s	.07	19	.03	9
	p × i	.01	3	.01	3
	s × i	.00	0	.00	0
	p × s × i, error	.03	8	.02	6

Notes. $N = 100$ for each instruction-type. $\delta^2(\bar{\alpha})$ = mean score estimated random effects variance component for $(\bar{\alpha})$. % var = percentage of variance accounted for (may not total 100% due to rounding error). Typical selection scenario = 1 person, 1 situation, 12 items.

cent of the total variance in personality responses across scales and instruction types.

Instruction type did seem to attenuate the situation and person × situation effects on personality responses. The average total percentage of variance accounted by the situation in the general instructions condition was approximately 14 percent whereas that same percentage in the applicant instructions condition was approximately 5 percent. Likewise,

the average total percentage of variance accounted by the person \times situation interaction in the general instructions condition was approximately 19 percent whereas that same percentage in the applicant instructions condition was approximately 10 percent. However, for three of the four factors, the person variance *increased* from the general instructions condition to the applicants instructions condition by an average of approximately 24 percent (whereas the person variance for Extraversion remained constant).

DISCUSSION

This partial reanalysis of the Schmit et al. (1995) data has provided additional insight into the effects of frame-of-reference on personality responses. The first hypothesis was fully supported; significant person variance was evidenced for every personality scale for both the general and applicant conditions. The second, third, and fourth hypotheses were partially supported. In regards to the second hypothesis, situation variance was not significant for the Extraversion scale in either the general or the applicant conditions, and situation variance was not significant for the Neuroticism and Agreeableness scales in the applicant condition. In regards to the third hypothesis, significant person \times situation variance was evidenced for every personality scale for both the general and applicant conditions and on average accounted for a higher percentage of the overall variance than the situation main effect; however, the relative sizes of these interactions did not conform exactly to our predictions. We hypothesized that the size of the person \times situation interactions would be highest for Neuroticism, followed by Conscientiousness, Agreeableness, and Extraversion, respectively; instead, we found the order to be: Neuroticism, Extraversion, Conscientiousness, and Agreeableness. We suspect that the lack of support for our prediction here may be due to sampling variability in the Extraversion, Conscientiousness, and Agreeableness estimates as they are all quite close to one another. Lastly, in regards to the fourth hypothesis, comparing the applicant to the general condition, person variance was *increased* in three of the four factors (Neuroticism, Agreeableness, and Conscientiousness), situation variance was reduced for the Neuroticism and Conscientiousness scales (with the situation variance being essentially the same between the two conditions for the Extraversion and Agreeableness scales), and person \times situation variance was reduced (but still statistically significant) for each scale.

Why was the person \times situation effect diminished in the applicant instructions condition? As noted earlier, two possible underlying mecha-

nisms are that: (1) the response distortion that is presumably infused when asking individuals to respond to a personality inventory as if they were applying for a job could act as error which may wash out any other experimental effects, or (2) when responding to a personality inventory as if one were applying for a job, individuals may already have an implicit work frame of reference which could act to negate any frame-of-reference person \times situation effect. As suggested earlier, the latter mechanism seems more plausible in that personality seems to evidence greater consistency and coherence when measured in similar situations (cf. Murtha et al., 1996); thus, the decrease in the person \times situation interaction may not be due to an increase in error associated with “faking” but to a decrease in non-relevant situational noise. The finding of interest was that, even with this clear press for adoption of a social role, the person \times situation interaction remained significant.

Why was the person variance larger in the applicant in comparison to the general conditions for the Neuroticism, Agreeableness, and Conscientiousness scales? This finding seems inconsistent with the findings from the Christiansen (1998) study that found the pre-post correlation between alternate measures of several personality constructs to be consistently lower for a group that was told to respond as if they were applying on the second administration in comparison to a group that was told to answer honestly on both administrations. Christiansen (1998) interpreted this finding as evidence that faking causes a reduction in “true variance.” However, Christiansen’s (1998) assumption that a low correlation between a general and applicant instructions condition is de facto evidence of a reduction in true variance is not in keeping with the findings of the present study. Our results suggest that Christiansen’s findings could perhaps be indicative of an *increase* in true variance because an increase in person variance in the current study would not have been evidenced if response distortion in the form of intentional faking were playing a major role. Consistent with the situational-dispositional representation of personality traits, when one is asked to respond to a personality measure for a given situation (e.g., applying for a sales position), situational “noise” is apt to be reduced and more relevant, “true” variance is apt to be increased (cf. Murtha et al., 1996).

In retrospect, it is perhaps not surprising that the Extraversion scale did not evidence an increase in person variance across the general condition as compared to the applicant condition given its genetic and physiological underpinnings. Specifically, personality factors that are more heavily influenced by genetic and physiological factors are probably less influenced by short-term environmental factors such as the manipulation we used in the present study. In regards to genetic factors, Loehlin (1992) conducted a meta-analysis which combined self-report data from

twin, adoption, and family studies into an extremely large data set to investigate the differential heritability of the Big Five. For all five factors, additive genetic effects were the strongest source of the phenotypic variance, explaining 43% of the variance in Culture/Openness, 32% of Extraversion, 27% of Neuroticism, 24% of Agreeableness, and 22% of Conscientiousness. A recent study specifically using NEO-FFI data found (for self-report) these corresponding estimates to be highest for Extraversion (56% of the variance) in relation to the other four factors (53%, 53%, 52%, and 42% for Openness/Culture, Conscientiousness, Neuroticism, and Agreeableness, respectively) (Riemann, Angleitner, & Strelau, 1997). In regards to physiological factors, unlike the other four factors, links have been made between individual differences in Extraversion and physiological etiology such as variations in baseline cortical activity (Eysenck, 1973; Gray, 1973).

The results of this partial reanalysis of Schmit et al.'s Study 1 data are consistent with findings from their second study. Specifically, they found that the criterion-related validity of several facets of the Conscientiousness construct were greater when contextualized items were used in an applicant instructions condition. A shift in rank order across situations (which is what a person \times situation interaction suggests) on the predictor side will, all else being equal, result in either an increase or decrease in validity across situations due to the nature of the correlation coefficient being affected by rank order. In the case of the facets of the Conscientiousness scale in Schmit et al.'s study, the situational manipulation (i.e., general vs. work-specific items) led to an increase in predictive validity—we might then expect that the person \times situation interaction evidenced in the applicant condition in the present study would likely suggest higher predictive validity for the work-specific items. Moreover, an increase in true variance (which is what the increase in person variance evidenced in the applicant condition in the present study suggests) will also likely lead to increases in predictive validity.

The present results then have possible implications for the use in personnel selection of the other three Big-Five personality factors that were included in our reanalysis of Schmit et al.'s study. Specifically, the person \times situation interactions for the Extraversion and Agreeableness scales were similar in magnitude to that of the Conscientiousness factor; however, that same effect for the Neuroticism factor was twice as large as the other three scales! Also, with the exception of the Extraversion scale, person variance increased from the general to the applicant instructions condition. In sum, our results suggest that, depending on the frame of reference, criterion-related validity and hiring decisions may be slightly different for Extraversion, larger still for Agreeableness and Conscientiousness, and markedly different for Neuroticism. If the effects

of frame-of-reference on criterion-related validity follow that of the Conscientiousness facets in Schmit et al.'s (1995) study, then the predictive validity of these Big-Five constructs may be significantly higher if properly contextualized in an applicant setting.

The implications of the present study's results for the use of the Neuroticism construct in personnel selection may be quite substantial. Although Barrick and Mount (1991) did not find Emotional Stability (reverse-coded Neuroticism) to be a good predictor of job performance across jobs and occupations ($\rho = .07$), two other meta-analyses by Tett, Jackson, and Rothstein (1991) and Salgado (1997) have found that Neuroticism ($\rho = -.22$) and Emotional Stability ($\rho = .18$), respectively had validities whose 90% credibility intervals did not encompass zero (i.e., it maintained at least a marginally useful level of validity). Regardless of these conflicting findings, strong evidence suggests that Neuroticism is related to measures of organizational delinquency (see Hough et al., p. 584). Moreover, research suggests that Neuroticism is a consistently found component of integrity tests—which are more and more being used for selection purposes (Hogan & Brinkmeyer, 1997; Ones, 1993). Thus, effects of frame-of-reference on the criterion-related validity of Neuroticism is highly consequential for personnel selection practice.

One criticism that can be leveled at the design of the present study is its mechanistic nature. Most interactional psychologists suggest that we should be studying interactions between persons and situations in a more organismic manner in which persons are perceived as active agents, influencing as well as being influenced by situations over a period of time (Endler & Magnusson, 1976; Hatstrup & Jackson, 1996). We agree with this notion; however, we believe that the setting to which we wish to generalize our results, a one-shot personnel selection scenario, justifies our use of the mechanistic model. A second limitation of the current study is that it is a laboratory study and generalizability of the results may be limited by this fact.

Future research needs to further establish the mechanisms behind the person \times situation interaction. For example, why is it that the rank order of persons shifts across frames-of-reference? Is personality at work a fundamentally different construct than personality in general? Future research should also further investigate what psychological components are contained in the person variance in general versus applicant conditions. Importantly, the question of whether increased person variance in applicant conditions harms or helps prediction of job-relevant outcomes should be answered, preferably using a between-groups, criterion-related, field study design. This research, and others of its type, should be replicated in applicant samples to strengthen arguments for generalizability.

REFERENCES

- Barrick, M. R., & Mount, M. K. (1991). The big five personality dimensions and job performance: A meta-analysis. *Personnel Psychology, 44*, 1–26.
- Bowers, S. K. (1973). Situationism in psychology: An analysis and a critique. *Psychological Review, 80*, 307–336.
- Christiansen, N. D. (1998, April). *Sensitive or senseless? Using social desirability measures to identify distortion*. Paper presented at the annual Society for Industrial and Organizational Psychology conference, Dallas, TX.
- Costa, P. T., Jr., & McCrae, R. R. (1989). *The NEO PI / FFI manual supplement*. Odessa, FL: Psychological Assessment Resources.
- Costa, P. T., Jr., & McCrae, R. R. (1992). *NEO PI-R professional manual*. Odessa, FL: Psychological Assessment Resources.
- Douglas, E. F., McDaniel, M. A., & Snell, A. F. (1996, August). *The validity of non-cognitive measures decays when applicants fake*. Paper presented at the annual Academy of Management conference, Cincinnati, OH.
- Endler, N. S., & Edwards, J. M. (1986). Interactionism in personality in the twentieth century. *Personality and Individual Differences, 7*, 379–384.
- Endler, N. S., & Magnusson, D. (Eds.) (1976). *Interactional psychology and personality*. Washington, DC: Hemisphere.
- Epstein, S., & O'Brien, E. J. (1985). The person-situation debate in historical and current perspective. *Psychological Bulletin, 98*, 513–537.
- Eysenck, H. J. (1973). *Eysenck on extraversion*. New York: Wiley.
- Frei, R. L., Griffith, R. L., Snell, A. F., McDaniel, M. A., & Douglas, E. F. (1997, April). Faking of non-cognitive measures: Factor invariance using multiple groups LISREL. In G. Alliger (Chair), *Faking matters*. Symposium conducted at the annual meeting of the Society for Industrial and Organizational Psychology, St. Louis, MO.
- Gray, J. A. (1973). Causal theories of personality and how to test them. In J. R. Joyce (Ed.), *Multivariate analysis and psychological theory* (pp. 409–464). New York: Academic Press.
- Hatrup, K., & Jackson, S. E. (1996). Learning about individual difference by taking situations seriously. In K. R. Murphy (Ed.), *Individual differences and behavior in organizations* (pp. 507–547). San Francisco, CA: Jossey-Bass.
- Hogan, R. (1991). Personality and personality measurement. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (Vol. 2, pp. 873–919). Palo Alto, CA: Consulting Psychologists Press.
- Hogan, J., & Brinkmeyer, K. (1997). Bridging the gap between overt and personality-based integrity tests. *Personnel Psychology, 50*, 587–599.
- Hough, L. M., Eaton, N. K., Dunnette, M. D., Kamp, J. D., & McCloy, R. A. (1990). Criterion-related validities of personality constructs and the effect of response distortion on those validities. *Journal of Applied Psychology, 75*, 581–595.
- Lakey, B., McCabe, K. M., Fisicaro, S. A., Drew, J. B. (1996). Environmental and personal determinants of support perceptions: Three generalizability studies. *Journal of Personality and Social Psychology, 70*, 1270–1280.
- Lindman, H. (1974). *Analysis of variance in complex experimental designs*. San Francisco: Freeman.
- Loehlin, J. C. (1992). *Genes and environment in personality development*. Newbury Park, CA: Sage.
- McCrae, R. R., & Costa, P. T., Jr. (1983). Social desirability scales: More substance than style. *Journal of Consulting and Clinical Psychology, 51*, 882–888.
- Mischel, W., & Shoda, Y. (1995). A cognitive-affective system theory of personality: Reconceptualizing situations, dispositions, dynamics, and invariance in personality structure. *Psychological Review, 102*, 246–268.
- Murtha, T. C., Kanfer, R., & Ackerman, P. L. (1996). Toward an interactionist taxonomy of personality and situations: An integrative situational-dispositional representation of personality traits. *Journal of Personality and Social Psychology, 71*, 193–207.
- Neter, J., Kutner, M. H., Nachtsheim, C. J., & Wasserman, W. (1996). *Applied linear statistical models* (4th ed.). Chicago, IL: Irwin.

- Olson, K. R., & Johnson, D. C. (1991). Individual differences in self-presentation styles. *Journal of Social Psychology, 131*, 495–509.
- Ones, D. S. (1993). *The construct validity of integrity tests*. Unpublished doctoral dissertation, University of Iowa, Iowa City, IA.
- Ones, D. S., Viswesvaran, C., & Reiss, A. D. (1996). Role of social desirability in personality testing for personnel selection: The red herring. *Journal of Applied Psychology, 81*, 660–679.
- Riemann, R., Angleitner, A., & Strelau, J. (1997). Genetic and environmental influences on personality: A study of twins reared together using the self- and peer report NEO-FFI scales. *Journal of Personality, 65*, 449–475.
- Roberts, B. W., & Donahue, E. M. (1994). One personality, multiple selves: Integrating personality and social roles. *Journal of Personality, 62*, 199–218.
- Salgado, J. F. (1997). The five factor model of personality and job performance in the European Community. *Journal of Applied Psychology, 82*, 30–43.
- Schmit, M. J., Ryan, A. M., Stierwalt, S. L., & Powell, A. B. (1995). Frame-of-reference effects on personality scale scores and criterion-related validity. *Journal of Applied Psychology, 80*, 607–620.
- Shavelson, R. J., & Webb, N. M. (1991). *Generalizability theory: A primer*. Newbury Park, CA: Sage.
- Snyder, M. (1983). The influence of individuals on situations: Implications for understanding the links between personality and social behavior. *Journal of Personality, 51*, 497–516.
- Tett, R. P., Jackson, D. N., & Rothstein, M. (1991). Personality measures as predictors of job performance: A meta-analytic review. *Personnel Psychology, 44*, 703–742.
- Wright, J. C., & Mischel, W. (1987). A conditional approach to dispositional constructs: The local predictability of social behavior. *Journal of Personality and Social Psychology, 53*, 1159–1177.
- Zickar, M. J. (1997, April). Computer simulation of faking on a personality test. In G. Alliger (Chair), *Faking matters*. Symposium conducted at the annual meeting of the Society for Industrial and Organizational Psychology, St. Louis, MO.