*MULTIVARIATE EFFECTS OF DEMAND
CHARACTERISTICS ON THE ANALOGUE ASSESSMENT
OF HETEROSOCIAL COMPETENCE*

JOSE A. MARTINEZ-DIAZ AND BARRY A. EDELSTEIN

WEST VIRGINIA UNIVERSITY

The present study assessed the effects of a contextual demand manipulation (pretreatment assessment *versus* experiment) and an instructional demand manipulation (high demand for competent behavior *versus* neutral demand) on the analogue assessment of heterosocial competence. Also assessed was the interaction between each demand manipulation and subject characteristics (high-frequency dating/low heterosocially anxious *versus* low-frequency dating/high heterosocially anxious college males). Students were assigned randomly to one of four groups receiving differential information and instructions. Students completed a battery of five standard paper-and-pencil questionnaires, responded out loud to 10 taped social situations, and conversed with a female confederate while being videotaped for five minutes. Multivariate analyses revealed highly significant differences between subject groups and between contextual manipulation groups. Univariate analyses revealed that only 3 of the 16 dependent measures, including 2 of the self-reports, were significantly affected by the contextual demand. Results are discussed in light of previous studies of phobic behavior and social skills. Various avenues for future research were presented.

DESCRIPTORS: analogue assessment, demand characteristics, multivariate effects, social skills

Several behavioral researchers have cautioned against the invalidating effect of demand characteristics on the assessment of clinically relevant behaviors in analogue settings (*e.g.*, Bernstein, 1973; Bernstein and Nietzel, 1977; Borkovec and O'Brien, 1976). There is a plethora of evidence that research in the area of phobic behaviors is plagued by the confounding effects of this artifact (*e.g.*, Bernstein and Nietzel, 1977; Borkovec, 1973, Note 1). Nietzel and Bernstein (1976, Note 2) report that assertive behaviors are affected by demand characteristics as well. In light of this evidence, it behooves clinical researchers to investigate the possible confounding influence of this "nuisance" variable (*cf.* Bernstein and Nietzel, 1977) on their assess-

ment procedures and on their outcome studies.

Even though analogue research on the assessment and training of heterosocial competence is currently receiving a great deal of attention (Hersen and Bellack, 1977; McFall, 1977), little has been done to study the effect of demand characteristics on this behavior class. The negative findings of the only study addressing this question (Borkovec, Stone, O'Brien, and Kaloupek, 1974) have been cited as evidence for the minimal need to control this artifact in research on heterosocial competence (*e.g.*, Hersen and Bellack, 1977). It is doubtful that the effects of demand characteristics on any behavior can be dismissed so easily.

The present study assessed the effects of demand characteristics on the analogue assessment of heterosocial competence. Following Cook and Campbell's (1975) recommendation that multiple operationalizations of an independent variable should be used, both a contextual and an instructional demand manipulation (*cf.* Bern-

This paper is based in part of the first author's master's thesis, directed by the second author. Many thanks to John D. Cone and Stan H. Cohen for their input into the master's thesis. Requests for reprints should be addressed to Barry A. Edelstein, Psychology Department, West Virginia University, Morgantown, West Virginia 26506.

stein, 1973) were used. The contextual manipulation consisted of varying the type of situation to which subjects were exposed. That is, subjects were randomly assigned to either an *experiment* or a pretreatment *assessment* context (Bernstein, 1973). The instructional manipulation consisted of randomly assigning subjects within each context to one of two different sets of instructions regarding how they were to behave during each of two tasks. That is, subjects were instructed to behave either in a *natural* or a *socially skillful, calm* manner. It was hypothesized that subjects in the assessment context would behave less skillfully and more anxiously than subjects in the experiment situation; that subjects instructed to behave skillfully and relaxed would behave more so than subjects instructed to behave naturally; and that the contextual manipulation would be more powerful than the instructional one.

The present study also assessed the interaction between subject characteristics (high-frequency dating/heterosocially relaxed *vs.* low-frequency dating/heterosocially anxious) and demand characteristics. Bernstein (1973) and Borkovec (1973) have hypothesized that the influence of demand characteristics may be attenuated by the selection of truly phobic subjects. A review by Borkovec (1973) of the selection criteria used in 18 different studies of the effects of demand characteristics on phobic behavior lend support to this hypothesis. The selection criteria used in studies reporting a significant difference between demand conditions were relatively weak, whereas studies reporting no demand effects used more stringent criteria (Borkovec, 1973).

The present study included two types of subjects: (1) subjects who self-reported both high-frequency dating and feeling calm and relaxed in heterosocial situations, and (2) subjects who self-reported both minimal dating and feeling anxious in heterosocial situations. Subjects within each type were randomly assigned to the four different conditions. It was hypothesized

that high-frequency dating/low heterosocially anxious subjects would be more susceptible to each demand manipulation than low-frequency dating/high heterosocially anxious subjects.

METHOD

Subjects

The Social Activity Questionnaire (Arkowitz, Lichtenstein, McGovern, and Hines, 1975) was administered to 876 undergraduate males enrolled at West Virginia University. The men taking this screening questionnaire were told that we were collecting normative data on the heterosocial behavior of college men and that their complete honesty would be appreciated. They were told, furthermore, that if they were interested in participating in one or more other projects dealing with heterosocial behavior they should write their name and telephone number; and that we would call most men who gave their telephone number. It was stressed that their answers to the questionnaire would not affect who was called. Students meeting one of the two following sets of criteria were telephoned and asked to participate in a study about college men's interactions with women: (1) students reporting four or more dates in the last month, 12 or more dates during the last six months, and feeling relaxed and comfortable in social situations with women were considered high-frequency dating/low heterosocially anxious men; (2) students reporting one or no dates during the last month, six or fewer dates during the last six months, and feeling "somewhat" or "very" anxious and inhibited in social situations with women were considered low-frequency dating/high heterosocially anxious men. In addition, only those students of either type who reported either a desire to date at least "somewhat" more than presently or a belief that their ability to make casual conversation with girls needed at least "some" improvement were telephoned. These potential subjects were asked to participate in one of several projects dealing

with male-female interactions. Regardless of project assignment, their participation would require 1 to 1½ hours and about half their time would be spent completing paper-and-pencil questionnaires. The remaining time would be spent responding out loud to a tape recorder and talking with a young woman. They were given no further information. If they asked any questions, they were told that further details depended on which project they participated in and that we did not assign subjects to the project until just before their participation in order to eliminate the possible selection bias. Indeed, this procedure was used to safeguard against selection bias. Of the 219 men who met the former criteria, 51 participated in the study. Of the 117 men who met the latter criteria, 42 participated. All men for whom telephone numbers were available were given the opportunity to participate. Thirty-five percent of the students who served as subjects received class bonus points; all others were volunteers. A preliminary MANOVA failed to show significant differences between subjects who participated for credit and those who volunteered, $F(15, 75) = 1.09, p > .05$. All dependent variables, except for heart rate change, were included in this MANOVA.

Settings and Materials

Two different experimental rooms were used. One was an office with a large desk where students completed self-report inventories. A tape recorder, placed on the desk, provided instructions. The second experimental room was electrically shielded and contained electrode leads, a video-camera, a tape recorder for instructions, two microphones, a table with physiological recording paraphernalia (e.g., electrodes, electrode paste), and two chairs. In addition to these two rooms, the experimenter occupied a room containing a Grass Model 78B polygraph, a Sony Model AV-3600 Videocorder and monitor, and a tape recorder for recording responses to a taped situation test (see below).

Design

A $2 \times 2 \times 2$ completely randomized mixed design (Neale and Liebert, 1973) was employed. Students from within each of two preselected dating frequency/heterosocial anxiety groups were assigned by way of a table of random numbers to one of four experimental groups. The two experimental variables manipulated to create the four groups were (1) situational context (experiment *vs.* pretreatment assessment) and (2) instructions (high *vs.* neutral demand for socially competent behavior).

Procedure

All students reported to the room containing the polygraph. After being informed of the general nature of their participation and signing consent forms, students were randomly assigned to one of the four conditions and began the study by either completing five questionnaires or participating in a taped situation test and *in vivo* interaction. Students first participating in the social interaction part of the experiment were seated in the shielded room containing the camera. Electrodes then were placed on the right breast, the lower left abdomen and the lower right abdomen. Students who first participated in the self-report portion of the study were seated at the desk in the other experimental room.

Students received their tape-recorded instructions while alone in each of the two experimental rooms. At the beginning of each of the two parts, students received information establishing their situational context. Situational context referred to the type of situation the student was led to believe he was in. Students in the *experiment* context ($n = 44$) were told that they were participating in an experiment and that their participation was a contribution to science. Those in the pretreatment assessment context ($n = 49$) were told that they were participating in an assessment conducted to determine if they qualified to participate in a train-

ing program designed to improve their ability to obtain dates and their heterosexual conversational skills. As a check on this manipulation, students were asked to describe the purpose of the study immediately after the first message establishing the context and again at the end of the experiment.

Students were asked to complete questionnaires, to respond out loud to each of 10 standard taped social situations, and to talk to a young woman they had met previously. The questionnaires used were chosen as the most commonly used in studies of heterosocial competence (Arkowitz, 1977). They included Watson and Friend's (1969) Social Avoidance and Distress scale and Fear of Negative Evaluation scale; Arkowitz, *et al.* (1975) modification of Endler, Hunt, and Rosenstein's (1962) Stimulus-Response Inventory of Anxiousness; Twen-tyman and McFall's (1975) Survey of Heterosexual Interactions; and Edwards' (1957) Social Desirability scale. The taped situations were the same as previously used by Arkowitz and his colleagues (1975). Immediately prior to the first taped situation, students were asked to behave either "in the *same way* you would if this were a *real* situation" (neutral instructional demand) or "in a *calm and relaxed* manner and in what you consider to be the *most socially skillful* way" (high instructional demand for socially skillful and relaxed behavior). Students were then asked to write down in what way we wanted them to respond to the taped situations as a manipulation check. Immediately before the extended interaction with the young woman, students were given the same instructions as above, but were not asked the manipulation check question. The *in vivo* interaction with the young woman lasted 5 min. It was purely conversational in nature, and was videotaped. Nine attractive college women served as confederates and were equally distributed across subject types and experimental conditions. Confederates were trained *via* behavior rehearsal and videotaped feedback to respond in a natural manner but not to initiate any topics of conver-

sation unless there was a silence of more than 15 sec. Confederates were naive as to the nature of the study and were ignorant of subject type and experimental conditions. They were kept in a room isolated from the experimental room until the time for interacting with subjects.

A multivariate analysis of variance failed to show significant differences due to confederates on measures taken from the *in vivo* conversation, $F(10, 79) = 1.02, p > .05$.

Measures

Besides the five standard questionnaires, the following measures were used to assess heterosocial competence: three measures were derived from the taped situation test: (1) an *average global rating of social skill* was obtained from observers' ratings of the student's audiotaped responses along a scale of 1 to 7 (1 = very socially unskilled; 7 = very socially skillful); (2) the *average number of words spoken* across situations was obtained from the same source; and (3) the *average latency of response* to the taped social situations was calculated. Seven measures were obtained from the videotaped extended interaction: (1) a *self-rating of social skill* was obtained from the student immediately after the confederate left the room using a scale identical to the one above; (2) a *self-rating of social anxiety* was obtained from the student at the same time, using a similar scale (1 = very anxious; 7 = very calm and relaxed); (3) an *average global rating of social skill* and (4) an *average global rating of social anxiety* were obtained from observers' ratings of the videotaped conversation, using scales identical to the ones above; (5) the *average amount of time spent gazing* at the confederate by the student; (6) the *average amount of time spent talking* by the student; and (7) the *average number of 10-sec silences* during the 5-min conversation were obtained from the videotape of the interaction. *Heart rate change* was obtained by randomly sampling six 10-sec periods from the polygraph record obtained during the taped situation test and six 10-sec samples from that obtained during

the *in vivo* interaction. The samples within each set were added to obtain heart rate in beats per minute. The score from the taped situation period was subtracted from the one for the conversation period to obtain a change score. This conversation estimate was used because it would not be as appropriate to use a period of rest as a baseline. Students were involved in similar activities both during the taped situation and the conversation, the only difference being the presence of the confederate in the latter one. That was hypothesized as the critical variable that might affect heart rate change. A discussion of the validity of the measures used is presented elsewhere (Martinez-Diaz and Edelstein, *in press*).

Observer agreement. Nine female research assistants, six of whom also served as confederates in the study, served as observers. They each rated the students' audio- and videotapes along global dimensions and recorded specific behaviors. The observers conducted their rating and recordings independently of each other. All observers were naive to the conditions and hypotheses of the study. Observers who also served as confederates were restricted from rating those subjects with whom they had interacted. Observers were restricted from rating students with whom they were acquainted. Each observer rated an equal number of students across subject types and experimental groups. (Between four and nine observers rated each student's tape for each of the variables.)

An average across observers was obtained for each of the global ratings by others and for each of the specific behaviors. Interobserver reliability was calculated for each possible pair of observers for each of the three global and each of the five specific behaviors rated from the video- and audiotapes. A mean reliability coefficient was obtained for each of these measures by using a transformation correcting for the effects of averaging correlations consisting of unequal *n*'s (McNemar, 1962, pp. 139-140) and using Spearman-Brown's formula which takes into account the number of judges used

(Winer, 1971, p. 288 ff). Reliability coefficients ranged from a low of .76 for the anxiety global rating to a high of .88 for both the number of 10-sec silences and the average number of words.

RESULTS

Check of Independent Variable Manipulations

Each of the two independent variable manipulations was checked twice; once immediately after the manipulation and once just before the student finished participation. Only if the student answered correctly both of the manipulation check questions relating to the particular independent variable was the identification considered correct. Even using this strict criterion, 98.9% of the students correctly identified their situational context; 75% of the students correctly identified their instructional demand. The student's ability to articulate the nature of the manipulation, however, was not necessarily assumed to be highly correlated with the extent to which the manipulation affected their behavior. The controversies notwithstanding, ample evidence exists in the literature for conditioning without "awareness".

A Priori Comparisons

Based on previous research in the area (*e.g.*, Bernstein, 1973), it was assumed that students would behave more similarly to the way they would in real-life situations in the experimental context and would behave less skillfully and more anxiously than in real-life situations in the pretreatment assessment. Given this assumption it was hypothesized that high-frequency dating/low heterosocially anxious students would be more likely to be affected by the contextual demand manipulation than the other students (Bernstein, 1973; Borkovec, 1973). It was also hypothesized that they would have more room to behave less skillfully and more anxiously in the pretreatment assessment context; that is, low frequency dating/high socially anxious students might experience a floor effect. Based on

these hypotheses, an *a priori* comparison including only high-frequency dating/low heterosocially anxious students was conducted. A Hotelling T^2 (Tatsuoka, 1971) was performed to test for differences between the experiment and assessment contexts for high-frequency dating/low heterosocially anxious men on all the dependent variables except for heart rate change. This *a priori* analysis yielded a nonsignificant T^2 of 29.82, $F(15, 35) = 1.42$, $p > .05$. However, an R^2 (Joe and Anderson, 1969) of .37 was derived from this analysis, indicating a moderate degree of association between the independent variable and a linear combination of dependent variables.

Multivariate Analyses of Variance

Because the contextual manipulation was aimed at all dependent variables, whereas the instructional manipulation was not aimed at the students' responses to the five standard questionnaires, analyses bearing on the effects of the latter independent variable excluded those dependent variables. Therefore, two separate multivariate analyses of variance (MANOVA's) were conducted. A two-way MANOVA tested for the effects of context and subject type on all dependent variables, except for heart rate change. Heart rate change was excluded in light of evidence supporting the law of initial values (Benjamin, 1963; Epstein, 1976; Johnson and Lubin, 1972; Lacey, 1956) which relates changes in heart rate to initial values on baseline. Heart rate change data were analyzed separately using an analysis of covariance (Benjamin, 1963). A second MANOVA (a three-way) was used to test for the effects of the instructional manipulation, as well as the other two independent variables, on all relevant dependent variables (*i.e.*, the same dependent variables used in the first MANOVA minus the five standard self-reports).

As mentioned above, a two-way (subject type $\times$ context) multivariate analysis of variance was conducted to test for significant differences between the two contexts, between the two sub-

ject types, and for the interaction between this independent variable and subject characteristics. This MANOVA yielded a significant main effect for the contextual manipulation, $F(15, 74) = 1.94$, $p < .05$. The strength of this association (1-Wilk's Lambda) was .28. The main effect of subject type was highly significant, $F(15, 74) = 3.65$, $p < .0001$. The measure of strength of association yielded a high .425. The interaction between subject type and context was nonsignificant, $F(15, 74) = 1.19$; whereas the strength of association between interaction and dependent variables was .194.

The three-way MANOVA testing for the main effects of the instrumental manipulation, the main effects of the contextual manipulation, the main effect of subject type, as well as their interactions, on the dependent variables derived from the taped situation test and the extended interaction, yielded the following results. There was no significant difference for the main effects of the instructional demand, $F(10, 75) = .034$, $p > .05$. The degree of association (1-Wilk's Lambda) between the 10 dependent variables and the instructional manipulation was .043. However, as in the above two-way MANOVA, both the main effects of the contextual manipulation, $F(10, 75) = 2.14$, $p < .05$, and of subject type, $F(10, 75) = 2.43$, $p < .01$, were found to be significant. The degree of association between this subset of the dependent measures and the above two independent variables was .222 and .245, respectively. There were no significant interaction effects (instructions $\times$ context, $F(10, 75) = .69$, $p > .05$, 1-Wilk's Lambda = .084; instructions $\times$ subject type, $F(10, 75) = .78$, $p > .05$, 1-Wilk's Lambda = .094; context by subject type, $F = .89$, $p > .05$, 1-Wilk's Lambda = .17; instructions $\times$ context $\times$ subject type, $F = .87$, $p > .05$, 1-Wilk's Lambda = .104).

Univariate Analyses of Variance

Because the multivariate F -test for the main effects of the contextual demand manipulation and of subject type proved significant in both the above MANOVAs, univariate F -tests were

calculated for each of the dependent variables to see which ones were significantly affected by these two independent variables (see Table 1). Instructional demand was excluded from these univariate analyses because the above multivariate F -tests testing for each effect proved to be nonsignificant. Only 3 of the 15 univariate F -tests yielded significant differences between the two contexts. Students in the experiment context scored significantly lower on the Fear of Negative Evaluation scale, $F(1, 88) = 3.85$, $p < .05$, and significantly higher on the Survey

of Heterosexual Interactions, $F(1, 88) = 4.58$, $p < .05$, than students in the pretreatment assessment context.

The former students also spent a significantly greater amount of time talking to the confederate than the latter students, $F(1, 88) = 4.51$, $p < .05$. However, estimates of Ω^2 for each of the above three analyses yielded values of only .024, .029, and .036, respectively. This indicates that the contextual manipulation accounted for only a small portion of the variance in these dependent variables. In contrast, 9 of

Table 1
Analysis of Variance

<i>Dependent Variable</i>	<i>Source</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>Estimate of Ω^2</i>
Social Anxiety and Distress Scale	SAQ-Type (A)	1	1055.79	32.27****	.253
	Context (B)	1	38.68	1.18	.001
	A $\times$ B	1	56.07	1.71	.006
	Within	88	32.72		
Fear of Negative Evaluation Scale	SAQ-Type (A)	1	1679.99	28.97****	.232
	Context (B)	1	223.23	3.85*	.024
	A $\times$ B	1	1.38	0.02	.000
	Within	88	1128.85		
Stimulus-Response Inventory of Anxiousness	SAQ-Type (A)	1	33813.41	29.95****	.238
	Context (B)	1	4025.45	3.57	.021
	A $\times$ B	1	20.26	0.02	.000
	Within	88	1128.85		
Survey of Heterosexual Interactions	SAQ-Type (A)	1	11101.25	29.11****	.236
	Context (B)	1	1745.11	4.58*	.029
	A $\times$ B	1	27.28	0.07	.000
	Within	88	381.33		
Edward's Social Desirability Scale	SAQ-Type (A)	1	328.92	14.24***	.124
	Context (B)	1	74.36	3.22	.021
	A $\times$ B	1	28.88	1.25	.002
	Within	88	23.10		
Self rating of social skill	SAQ-Type (A)	1	25.64	14.39***	.128
	Context (B)	1	2.07	1.16	.002
	A $\times$ B	1	1.18	0.66	.000
	Within	88	1.78		
Self rating of social anxiety	SAQ-Type (A)	1	16.94	10.25**	.091
	Context (B)	1	3.02	1.83	.008
	A $\times$ B	1	2.12	1.28	.003
	Within	88	1.65		
Video rating by other of social skill	SAQ-Type (A)	1	3.25	3.32	.024
	Context (B)	1	2.28	2.33	.014
	A $\times$ B	1	0.68	0.69	.000
	Within	88	0.98		

(continued)

Table 1 *continued.*

<i>Dependent Variable</i>	<i>Source</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>Estimate of Omega²</i>
Video rating by other of social anxiety	SAQ-Type (A)	1	6.58	6.49**	.057
	Context (B)	1	0.23	0.23	.000
	A × B	1	0.73	0.72	.000
	Within	88	1.01		
Video time spent talking	SAQ-Type (A)	1	3268.72	3.81*	.029
	Context (B)	1	3862.61	4.51*	.036
	A × B	1	1.91	0.00	.000
	Within	88	857.18		
Video time spent gazing	SAQ-Type (A)	1	487.84	0.22	.000
	Context (B)	1	1011.47	0.45	.000
	A × B	1	3579.48	1.59	.006
	Within	88	2250.61		
Video number of 10-sec silences	SAQ-Type (A)	1	0.007	0.09	.000
	Context (B)	1	0.003	0.03	.000
	A × B	1	0.367	4.73*	.040
	Within	88	0.078		
Average number of words	SAQ-Type (A)	1	16.58	0.39	.000
	Context (B)	1	120.94	2.88	.020
	A × B	1	55.24	1.31	.003
	Within	88	42.05		
Audio latency of response	SAQ-Type (A)	1	0.005	0.01	.000
	Context (B)	1	1.546	2.70	.018
	A × B	1	1.896	3.31	.025
	Within	88	0.573		
Audio rating by other of social skill	SAQ-Type (A)	1	0.08	0.08	.000
	Context (B)	1	0.09	0.08	.000
	A × B	1	3.48	3.18	.024
	Within	88	1.09		

**** $p < .0001$.*** $p < .001$.** $p < .01$.* $p < .05$.

the 15 univariate tests were significant for subject type (see Table 1). All these differences were in the expected direction (*cf.* Martinez-Diaz and Edelstein, *in press*). Estimates of Ω^2 for subject type were generally higher than values obtained for the contextual manipulation.

Analysis of Covariance (Heart Rate Change Data)

As mentioned above, the effects of the independent variables on heart rate change were analyzed separately because it was necessary to partial out effects of the law of initial values

(Benjamin, 1963). A three-way analysis of covariance was conducted, using heart rate baseline as a covariance, to test for the effects of the two independent variables (Benjamin, 1963). This analysis resulted in no significant main effects for subject type, $F(1, 80) = 3.54$, $p > .05$, context, $F(1, 80) = 2.61$, $p > .05$, and instructions, $F(1, 80) = 0.54$, $p > .05$. A significant interaction for subject type by instructions was found, $F(1, 80) = 6.80$, $p < .01$, but not for any other interaction; whereas high-frequency dating/low heterosocially anxious students in the neutral instruction condition decreased their heart rate from baseline to the *in vivo* interac-

tion; the students of this type in the high demand instruction condition increased their heart rate. The opposite was true for low-frequency dating/high heterosocially anxious students. Significant differences for the effects of the covariate were found, $F(1, 80) = 7.32, p < .01$.

DISCUSSION

The effects of the demand manipulations were equivocal. As hypothesized, the contextual manipulation was more potent than the instructional one. Students behaved more similarly to the manner in which they would in real-life situations in the experimental context and behaved less skillfully and more anxiously than in real-life situations in the pretreatment assessment. Thus, situational context is apparently a good predictor of heterosocial behavior, particularly for individuals who report relatively high frequency dating and low heterosocial anxiety; the degree of association between situational context and the dependent variables being .37 for this subject type. A particularly important finding was that the effects of the contextual manipulation were detected only by multivariate data analysis techniques. In light of these findings, the question arises whether the lack of demand effects on heterosocial behavior reported by Borkovec and his colleagues (1974) may be due to the insensitivity of their univariate analyses to trends in the data. The instructional manipulation had little or no effect on behavior. When students of both types and in both situational contexts are considered, instructional demand was a poor predictor of behavior. One may tentatively conclude, however, that instructional demand is a fairly good predictor of heterosocial competence when the subject sample is relatively homogeneous and when the context is experimental. It is a poor predictor when the sample is relatively heterogeneous and both the experiment and pretreatment assessment contexts are used. The generally mixed effects of instructional demand are concordant with the finding of Borkovec *et al.* (1974). In contrast,

Nietzel and Bernstein (1976, Note 2) have reported strong instructional demand effects on assertive behavior, a related behavior class. One plausible *post hoc* explanation for the divergence of the two sets of results is that Nietzel and Bernstein informed their subjects about the general type of responses that would be considered assertive and about some procedures for their assessment (*cf.* Nietzel and Bernstein, Note 2, p. 4.). Berkovec *et al.* (1974) and the present authors did not provide this information. Thus demand characteristics may have affected primarily those behaviors which the subject was aware were being measured and/or which the subject was aware constituted an expected way of responding (Orne, 1969). For example, demand characteristics may affect approach behavior for phobic subjects in a phobic situation since the subject may clearly see what is being measured by a Behavior Avoidance Test (BAT) and the meaning of a particular way of responding. Some degree of corroboration for this hypothesis is provided by the fact that the Survey of Heterosexual Interactions, the most face valid measure of heterosocial competence, was the dependent variable most strongly affected by the contextual manipulation.

The MANOVA's and univariate analyses clearly indicated that subject type accounted for a much greater percentage of the variance in the dependent variables than either of the demand manipulations (Martinez-Diaz and Edelstein, *in press*). This finding may be explained by the commonality in method between the present subject (type) selection criteria and 7 of the 16 measures (Bellack and Hersen, 1977; Cone, 1977; Martinez-Diaz and Edelstein, *in press*). That is, both the Social Activity Questionnaire (the selection instrument) and seven dependent variables are self-reports. The pervasiveness of method variance as a powerful factor in assessment may be inferred from the work of several researchers (*e.g.*, Campbell and Fiske, 1959; Mischel, 1968). Examination of assessment studies in the area (*e.g.*, Arkowitz *et al.*, 1975; Glasgow and Arkowitz, 1975; Twen-

tyman and McFall, 1975) also support the assertion that method commonality accounts for the largest share of variance in the assessment of heterosocial competence.

There was no evidence of either a subject characteristics by contextual demand interaction or a subject characteristics by instructional demand interaction, which is contrary to hypotheses of Bernstein (1973), Bernstein and Nietzel (1977), Borkovec (1973), and Borkovec and O'Brien (1976). Previous research (*e.g.*, Bernstein and Nietzel, 1977, pp. 135-138; Borkovec, 1973, pp. 493-495) using direct observations and self-reports would suggest that the low-frequency dating/high heterosocially anxious subjects in the present study were too skilled and relaxed to be differentially affected by the demand characteristics. It is also possible that students behaved in a way consistent with their preestablished social role during the study, being more influenced by implicit demands than by the manipulated demands. Subjects may have behaved consistently with the demands established prior to the study while completing the Social Activity Questionnaire.

The foregoing evidence leads to the conclusion that the issue of whether the assessment of heterosocial competence is affected by demand characteristics is far from settled. These data also lend support to multivariate approaches in the evaluation of data derived from complex combinations of independent and dependent variables. The use of such techniques in the behavioral literature is almost nonexistent and is certainly not incompatible with subsequent single-subject research which manipulates those variables accounting for the most variance in predicting behavior.

Future research should address at least the following: (1) the effects of demand characteristics on variations of assessment procedures and (2) the different types of demand characteristics which may affect assessment procedures. This research will hopefully lead to the control of this "artifact" and its utilization in the assess-

ment of heterosocial competence (Bernstein and Nietzel, 1977).

REFERENCE NOTES

1. Borkovec, T. D. Expectancy: Demand for improvement and/or cognitive set. Paper presented at the ninth annual meeting of the Association for Advancement of Behavior Therapy, San Francisco, December, 1975.
2. Nietzel, M. T. and Bernstein, D. A. *The effects of instructionally-mediated demand upon the behavioral assessment of assertiveness*. Unpublished manuscript, University of Kentucky, 1975. (Available from Michael T. Nietzel, Department of Psychology, University of Kentucky, Lexington, Kentucky 40506.)

REFERENCES

- Arkowitz, H. The measurement and modification of minimal dating behavior. In M. Hersen, R. Eisler, and P. Miller (Eds), *Progress in behavior modification*, Vol. 5. New York: Academic Press, 1977.
- Arkowitz, H., Lichtenstein, E., McGovern, K., and Hines, P. The behavioral assessment of social competence in males. *Behavior Therapy*, 1975, 6, 3-13.
- Bellack, A. S. and Hersen, M. Self-report inventories in behavioral assessment. In J. D. Cone and R. P. Hawkins (Eds), *Behavioral assessment: New directions in clinical psychology*, Brunner/Mazel: New York, 1977.
- Benjamin, L. S. Statistical treatment of the law of initial values (LIV) in autonomic research: A review and a recommendation. *Psychosomatic Medicine*, 1963, 25, 556-566.
- Bernstein, D. A. Behavioral fear assessment: Anxiety or artifact? In H. Adams and P. Unikel (Eds), *Issues and trends in behavior therapy*, Springfield, Illinois: Charles C Thomas, 1973.
- Bernstein, D. A. and Nietzel, M. T. Demand characteristics in behavior modification: The natural history of a "nuisance." In M. Hersen, R. Eisler, and P. Miller (Eds), *Progress in behavior modification*, Vol. 4. New York: Academic Press, 1977.
- Borkovec, T. D. The role of expectancy and physiological feedback in fear research: A review with special reference to subject characteristics. *Behavior Therapy*, 1973, 4, 491-505.
- Borkovec, T. D. and O'Brien, G. T. Methodological and target behavior issues in analogue therapy outcome research. In *Progress in behavior modification*, Vol. 3. New York: Academic Press, 1976.

- Borkovec, T. D., Stone, N. M., O'Brien, G. T., and Kaloupek, D. G. Evaluation of a clinically relevant target behavior for analogue outcome research. *Behavior Therapy*, 1974, **5**, 504-514.
- Campbell, D. T. and Fiske, D. W. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 1959, **14**, 81-105.
- Cone, J. D. The relevance of reliability and validity for behavioral assessment. *Behavior Therapy*, 1977, **8**, 506-520.
- Cook, T. D. and Campbell, D. T. The design and conduct of quasiexperiments and true experiments in field settings. In M. D. Munnette (Ed), *Handbook of industrial and organizational research*. New York: Rand McNally, 1975.
- Edwards, A. L. *The social desirability variable in personality assessment and research*. New York: Holt, Rinehart & Winston, 1957.
- Endler, W. S., Hunt, J. McV., and Rosenstein, A. J. An S-R Inventory of anxiousness. *Psychological Monographs*, 1962, **76** (17, Whole No. 536).
- Epstein, L. H. Psychophysiological measurement in assessment. In M. Hersen and A. S. Bellack (Eds), *Behavioral assessment: A practical handbook*, Elmsford, New York: Pergamon, 1976.
- Glasgow, R. E. and Arkowitz, H. The behavioral assessment of male and female social competence in dyadic heterosexual interactions. *Behavior Therapy*, 1975, **6**, 488-498.
- Hersen, M. and Bellack, A. S. Assessment of social skills. In A. R. Ciminero, K. S. Calhoun, and H. E. Adams (Eds), *Handbook for behavioral assessment*, New York: Wiley, 1977.
- Joe, G. W. and Anderson, H. E. Mean differences and association. *Multivariate Behavioral Research*, 1969, **4**, 379-389.
- Johnson, L. C. and Lubin, A. On planning psychophysiological experiments: Design, measurement and analysis. In N. S. Greenfield and R. A. Sternbach (Eds), *Handbook of psychophysiology*. New York: Holt, Rinehart & Winston, 1972.
- Lacey, P. J. The evaluation of autonomic responses: Toward a general solution. *Annals of the New York Academy of Science*, 1956, **67**, 123-164.
- Martinez-Diaz, J. A. and Edelstein, B. Heterosocial competence; Predictive and construct validity. *Behavior Modification*, in press.
- McFall, R. M. Analogue methods in behavioral assessment: Issues and prospects. In J. B. Cone and R. P. Hawkins (Eds), *Behavioral assessment: New directions in clinical psychology*. New York: Brunner/Mazel, 1977.
- McNemar, Q. *Psychological statistics* (3rd ed.). New York: Wiley 1962.
- Mischel, W. *Personality and assessment*. New York: Wiley, 1968.
- Neale, J. and Liebert, R. *Science and behavior: An introduction to methods of research*. Englewood Cliffs, New Jersey: Prentice Hall, 1973.
- Nietzel, M. T. and Bernstein, D. A. Effects of instructionally mediated demand on the behavioral assessment of assertiveness. *Journal of Consulting and Clinical Psychology*, 1976, **44**, 500.
- Orne, M. T. Demand characteristics and the concept of quasi-controls. In R. Rosenthal and R. L. Rosnow (Eds), *Artifact in behavioral research*. New York: Academic Press, 1969.
- Tatsuoka, M. M. *Multivariate analysis: Techniques for educational and psychological research*. New York: Wiley, 1971.
- Twentyman, C. T. and McFall, R. M. Behavioral training of social skills in shy males. *Journal of Consulting and Clinical Psychology*, 1975, **43**, 384-395.
- Watson, D. and Friend, R. Measurement of social-evaluative anxiety. *Journal of Consulting and Clinical Psychology*, 1969, **33**, 448-457.
- Winer, B. J. *Statistical principles in experimental design*. New York: McGraw-Hill, 1971.

Received 28 September 1978.

(Final Acceptance 31 August 1979.)