CURRENT RESEARCH IN SOCIAL PSYCHOLOGY

http://www.uiowa.edu/~grpproc/crisp/crisp.html

Volume 13, No. 15
Submitted: February 27, 2008
First Revision: March 19, 2008
Accepted: June 24, 2008
Published: July 13, 2008

REPORTING DIFFERENCES IN DEPRESSION AND HEALTH RISK EVALUATIONS IN DISSIMILAR TESTING CONDITIONS

Sarah-Jane Renaud
McGill University

Stewart Page
University of Windsor

ABSTRACT

Following earlier research concerning the effects of situational cues and social role demands on test item responses, the present study examined the effects of differences in the surrounding social situation, and in the ostensible purpose of testing, as determinants of test item responses and related health questions. Participants were 143, 34 male and 109 female, undergraduate students who responded to the Beck Depression Inventory, plus twelve items assessing various somatic complaints from the Hopkins Symptom Checklist, five items assessing concern for selected health problems, and three questions measuring levels of concern regarding two written health risk scenarios. Participants were randomly assigned to either a clinically "serious" or informally "casual" testing condition. Univariate analyses showed that the main effects of participant gender and condition on some measures were nonsignificant. A multivariate analysis of variance, however, including all dependent variable measures together, showed the main effect of testing condition to be significant, $F(9, 126) = 2.52$, $p < .023$. In our opinion, the effects upon test responses of surrounding situational cues, including the related issue of gender, are still a live issue in the social psychology of test-taking and research situations in general.

INTRODUCTION

In an important study of the social psychology of psychological testing, Kroger and Turnbull (1970) (see also Earn & Kroger, 1976; Page & Yates, 1975) found large and significant differences in responses to test items when participants were tested by a uniformed military officer, in a military building, with the ostensible purpose of testing being the assessment of "officer effectiveness," as opposed to when comparable participants completed identical tests, portrayed as assessing "artistic ability," in an artist's studio within a fine arts building, overseen by an ostensible professor of art.

More recently, the issue of role demands created by cues within the surrounding testing situation has been related to the areas of adjustment and depression and their assessment by means of psychological testing (e.g., Goodwin & Gotlib, 2004; Page, 1999; Page & Bennesch, 1993; Vredenburg, Krames & Flett, 1986). The possibility that gender differences in depression are related, for example, to males concealing or modifying their acknowledgement of depression on tests, as found, for example, by Vredenburg, Krames, & Flett (1986), also remains largely unaddressed.

Page and Bennesch (1993) examined whether the explicit portrayal of a test as a measure of depression affected the level of endorsement of items on the Beck Depression Inventory (BDI) and the Hopkins Symptom Checklist (HSCL). These authors employed a condition in which participants answered BDI items, described explicitly as assessing depression, and as ostensibly forming part of a Depression Research Project. In another condition, participants responded to the identical items, but with an accompanying procedure which made no mention of depression, and in which item content was described as assessing the "hassles of everyday life." Male participants, though not females, generally endorsed significantly more depression and adjustment items when they believed the items assessed only "hassles."

The present study similarly examined whether differences in testing conditions would affect test-taking. It was hypothesized that participants, perhaps especially males, would be more likely to acknowledge behaviors or symptoms when the surrounding situation was casual in tone, but that their responses would be more constrained when the surrounding situation was serious in tone. Accordingly, participants completed BDI and HSCL items in two testing conditions, varying in cues connoting ostensible "clinical seriousness," that is, in their formality and level of apparent credulity toward the seriousness of clinical symptoms.

Participants also responded to several items concerning their level of concern (that is, after considering them in the context of their family history and current lifestyle) for five major health problems (i.e., heart attack, stroke, diabetes, skin and lung cancer), and concerning their reactions to two written health risk scenarios, similar to those used by Lafreniere, Out, & Cramer (2006).

One written scenario described the occurrence of developing diabetes. The other scenario described the development of brain cancer from cellular phone use, in order to assess responses regarding a contentious, yet largely unsubstantiated disorder (National Cancer Institute, 2004; Nordenberg, 2000).

For each of the written scenarios, participants indicated on five-point Likert scales their level of concern for development of the disorder, the perceived likelihood of the disorder, and the level of action believed necessary to avoid the risk described. Following the above, the hypothesis was examined that ratings of concern, and level of action perceived as merited, would be more constrained in the serious, relative to the casual, testing situation.

METHOD

Participants

Participants were 143, 34 male and 109 female, undergraduate students at the University of Windsor, recruited through the online departmental participant pool. Course credits were given for participation.

All participants signed a written consent form, submitted separately from the individual's data. Upon completion, participants were given a brief verbal description of the study and information with which they could request a full report of the study.

Materials

All participants completed a questionnaire, composed of the revised 21-item BDI (Beck, Rush, Shaw, & Emery, 1979) and twelve items (nos. 1, 4, 12, 14, 27, 42, 48, 49, 52, 53, 56, and 58) assessing complaints related to depression, taken from the HSCL (Derogatis, Lipman, Rickels, Uhlenruth, & Covi, 1974). The BDI was chosen to enable comparisons with recently published research concerning the issue of test-taking differences as affected by situational cues.

After being asked to consider the context of their personal lifestyle and family history, participants were then asked to rate their degree of concern for the five health problems, that is, skin and lung cancer, stroke, heart attack, and diabetes, and to provide these responses along a four-point Likert scale ranging from 1 (not concerned) to 4 (very concerned).

The two health risk scenarios were then presented. Using five-point Likert scales, participants indicated their level of concern for the health issue described, its perceived likelihood of occurrence, and the level of action they believed they would take to avoid the issue described.

Design and Procedure

In the "clinical" condition, a face sheet described the ostensible purpose of the study. An attempt was made to replicate the quiet atmosphere of a medical waiting room. The tester affected a friendly but serious demeanor, wore a white lab coat, read the testing instructions in a formal manner, and directed participants as to where to be seated. The labels "depression inventory" and "medical symptoms" appeared on the written materials, and the study was described as concerning the identification of depressive symptoms in the current population, with the study's questionnaire portrayed as a potential assessment tool in medical research or clinical practice. On each page, under the title "Illness Evaluation," was the phrase "To be administered prior to initial

medical evaluation." At the bottom of each page, a footer read "Medical Health Research Team, 2006." When participants arrived, the researcher (SJR) followed a formal script including (a) asking the participant his or her name, (b) asking the participant to leave belongings at the door and to quietly take a specific seat, and (c) giving instructions on how to exit the room. When all participants had been seated, the consent forms were completed, the researcher read aloud the purpose of the study, and then administered the measures described above.

In the "casual" condition, the student-based cafeteria at the University of Windsor constituted the testing situation. Participants were met informally by the tester, asked to "just grab a seat anywhere" before "handling some questions," and were thus fully exposed to other students socializing, eating, and studying, plus the noise and other "stimuli" of a student cafeteria. The researcher (SJR) read the testing instructions in a casual fashion. No "serious" or specific purpose of the study was conveyed, all titles and special terms were removed from the questionnaire, and the general approach of the tester was informal, casual, and friendly. In order to maintain the informal nature of the condition, participants in this condition did not complete the consent form until after completion of the questionnaire, but initially were informed that they were under no obligation to answer all of the questions and could refrain from participation at any time. Further, data-gathering was described as being concerned only with assessing the "hassles of everyday life."

RESULTS

Tables 1 and 2 show the means, standard deviations, and group sizes for each of the dependent variables. BDI and HSCL scales were significantly correlated, at $r = .218$, $p < .01$.

Table 1. Overall BDI and HSCL Means and Standard Deviations by Condition and Gender

Measure	Clinical M	Casual M	Clinical SD	Casual SD	Clinical n	Casual n
BDI	8.69	11.41	7.03	13.11	65	77
Males	9.75	13.36	10.70	20.79	8	25
Fem.	8.54	10.45	6.48	6.91	57	51
HSCL	8.37	9.81	5.71	6.01	63	77
Males	8.14	10.04	7.49	5.71	7	26
Females	8.39	9.69	5.54	6.20	56	51

Table 2. Health Risk Scenario Means and Standard Deviations for Diabetes and Brain Cancer From Cellular Phones, by Gender and Condition

Measure	Clinical M	Casual M	Clinical SD	Casual SD	Clinical n	Casual n
ConcernDiab.	3.76	3.68	1.12	1.08	66	77
Males	3.63	3.38	1.51	1.27	8	26
Females	3.78	3.82	1.08	.95	58	51
LikelihDiab	3.48	3.36	.86	1.01	66	77
Males	3.25	2.92	1.17	1.09	8	26
Females	3.52	3.59	.82	.90	58	51
ActionDiab	3.73	3.90	1.24	1.11	66	77
Males	3.25	3.85	1.91	1.26	8	26
Females	3.79	3.92	1.12	1.04	58	51
ConcernCanc	2.45	2.55	1.15	1.25	66	77
Males	2.63	2.58	1.77	1.45	8	26
Females	2.43	2.53	1.06	1.15	58	51
LikehiCanc	2.03	2.26	.87	.95	66	77
Males	1.63	2.50	.74	1.14	8	26
Females	2.09	2.14	.88	.83	58	51
Cell Action	3.52	3.95	1.54	1.24	66	77
Males	2.88	4.12	1.89	1.21	8	26
Females	3.60	3.86	1.49	1.27	58	5

A 2 (Subject Gender) x 2 (Condition) MANOVA, with only the BDI and HSCL scores as dependent variables, found that the main effects for gender and condition, and interaction, were nonsignificant.

When all dependent variables were included, that is, scores on the BDI, HSCL, the two health risk scenarios, and scores on the five health problems, a MANOVA by gender and testing condition was significant for the effect of testing condition, F (9, 126) = 2.52, $p < .023$.

Univariate analyses

Separate ANOVAs univariate analyses of variance (ANOVAs), by gender and testing condition, for the BDI and HSCL scales, were both nonsignificant.

Individual 2 x 2 ANOVAs were performed for each of the five health problem evaluations. Although the effects regarding perceived risk of diabetes, stroke, and heart attack, were nonsignificant, an ANOVA regarding concern for lung cancer showed that the effect of gender was significant, F (1, 139) = 4.006, $p < .047$, with males showing higher ratings of concern for this problem.

Regarding concern for skin cancer, the effect for gender was significant, F (1, 139) = 4.854, p < .029, with females reporting higher levels of concern, relative to males. The interaction concerning skin cancer was also significant, i.e., at F (1, 139) = 3.836, p < .052, that is, males reported far less concern in the clinical condition, relative to females and to the level reported by males in the casual condition. In the casual condition, females reported greater concern for skin cancer risk, relative to males, and less perceived risk compared to females in the clinical condition.

Diabetes Health Risk Scenario

A 2 x 2 ANOVA, gender by testing condition, was conducted for each of the three questions asked after reading of the diabetes health risk scenario. Although the interaction and main effects for level of concern and estimated level of action contemplated were nonsignificant, the effect for gender concerning the perceived likelihood of diabetes was significant, F (1, 139) = 5.105, p < .025. Females thus rated the perceived likelihood of diabetes higher than did males.

Brain Cancer Health Risk Scenario

For the brain cancer scenario, a 2 x 2 ANOVA found that the effect for testing condition was significant for ratings of perceived likelihood of this problem, F (1, 139) = 5.835, p < .017, and for the reported level of action a participant would take to avoid this risk, F (1, 139) = 5.175, p < .024. Individuals in the casual condition generally expressed greater concern, yet also indicated they were likely to use their cellular phone as their primary telephone, notwithstanding their perception of elevated risk in so doing, relative to those in the clinical condition. The interaction concerning the level of action participants would take to avoid brain cancer risk was significant, F (1, 139) = 4.097, p < .045. Females reported nearly the same level in both conditions; males in the casual condition, however, were more likely to indicate cellular phones as their primary phone.

DISCUSSION

The hypothesis that BDI and HSCL scores would differ by testing condition was unsupported by some of the ANOVA results. Mean BDI scores, by condition, were however nearly three points apart, with the casual condition more likely to elicit symptoms of depression or maladjustment. The mean BDI scores in the casual condition were just above the range (i.e., 10-11) often considered as indicating mild depression (Goodwin & Gotlib, 2004).

The general hypothesis that concern for health problems would be related to differences in testing situation was not uniformly supported, although the main effects for gender were significant for evaluations of skin and lung cancer risk, with males showing greater concern for lung and females greater concern for skin cancer.

The general hypothesis of testing situation differences regarding ratings of concern, likelihood, and level of action in regard to the two written health risk scenarios, did receive support concerning the scenario describing risk of brain cancer from cellular phones. The casual condition elicited significantly higher scores, showing higher levels of concern relative to the

clinical condition, with males in the casual condition especially likely to show greater concern. Participants in the casual condition perceived considerable risk, yet were more likely, however, to describe cellular phones as their primary phone.

We note that, perhaps contrary to expectation, gender and depression scores in the present data were uncorrelated.

We recognize that the results were uneven and did not reach statistical significance in every comparison. Moreover, the number of male participants ($N = 34$) in our sample afforded less generalizeability and reduced power for some comparisons. The preponderance of female participants is, however, characteristic of both the University of Windsor participant pool and that of many Canadian universities, that is, with far fewer male students registered in psychology courses.

In general, when considering all of our dependent measures together as measures of depression and related symptoms, participants endorsed these at higher levels when the surrounding situation was casual. By contrast, responses were generally more constrained when testing, using identical test items, occurred in a more formal situation. As a researchable hypothesis for the future, we would speculate therefore that casual or "peer-oriented" testing situations may elicit more candid, less guarded, and thus more accurate, responding in terms of acknowledgment of maladjustment. In related studies, Page and Bennesch (1993), also Page (1999), found that participants, particularly males, were more likely to endorse items assessing depression when these were presented casually as assessing the "hassles of everyday life," compared to when the identical items were portrayed as formal symptoms of depression. The seemingly greater willingness of females to acknowledge or "admit" depression or other health concerns in more formal situations, together with the lower likehihood of males to acknowledge them, seems consistent with common male and female gender role stereotypes. In the present data, nonetheless, both males and females were less likely to acknowledge maladjustment in the clinical situation.

In our opinion, variables such as the surrounding testing situation must be considered in interpretation of test-taking responses. This notion follows the early writing of Campbell and Fiske (1959), who conceptualized test scores as interactive products of both underlying traits and methods for their assessment. To the parameters of trait and method, we would thus add the characteristics of the testing situation, including the potential main or interactional effects of factors such as gender and the ostensible purpose of testing to which participants are sensitized and observed to react. The notion of reactivity of data gathering to surrounding cues was raised principally by Webb, et. al (1966). Although no longer "in vogue," the area known as the social psychology of the psychological experiment, initiated largely by Orne (1962) and Rosenthal (1976), should in our view be revisited, out of concern with situational effects upon test responses and upon the behavior of research participants in general.

REFERENCES

Beck, A., Rush, A., Shaw, B., & Emery, G. (1979). *Cognitive therapy of depression*. New York: Guilford Press.

Campbell, D., & Fiske, D. Convergent and discriminant validation by the multitrait multimethod matrix. *Psychological Bulletin, 56*, 81-105.

Derogatis, L. R., Lipman, R., Rickels, K., Uhlenhuth, E. H., & Covi, L. (1974). The Hopkins symptom checklist (HSCL): A self-report symptom inventory. *Behavioural Science, 19*, 1-15.

Earn, B. M. & Kroger, R. O. (1976). The subject in psychological experiments: Effects of experimentally induced roles on laboratory performance. *Personality and Social Psychology Bulletin, 2(4*), 466-469.

Goodwin, R. D. & Gotlib, I. H. (2004). Gender differences in depression: the role of personality factors. [Electronic Version]. *Psychiatry Research, 1269(2),* 135-142.

Katz, R. & McGuffin, P. (1987). Neuroticism in familial depression. *Psychological Medicine, 17(1)*, 155-161.

Kroger, R. O. (1967). Effects of role demands and test-cue properties upon personality test performance. *Journal of Consulting Psychology, 31(3*), 304-312.

Kroger, R. O. & Turnbull, W. (1970). Effects of role demands and test-cue properties on personality test performance: Replication and extension. *Journal of Consulting and Clinical Psychology, 35(3*), 381-387.

Lafreniere, K., Out, J. & Cramer, K. (2006). *Risk perception and concern about probable and improbable health risks: A reversal theory approach*. Paper presented at the Annual Meeting of the Canadian Psychological Association, Calgary, Alberta.

Miles, A., Waller, J., Hiom, S., & Swanston D. (2005). SunSmart? Skin cancer knowledge and preventive behaviour in a British population representative sample. *Health Education Research, 20(5)*, 579-585.

NCI health information tip sheet for writers: Cellular telephones and cancer. (2004, August 16). *National Cancer Institute*. Retrieved December 10, 2005 from *http://www.cancer.gov/newscenter/tip-sheet-cellular-telephones.html*

Nordenberg, T. (2000, December). Cell phones and cancer: No clear connection. *Food and Drug Administration Consumer Magazine, 6*. Retrieved December 10, 2005 from *http://www.fda.gov/fdac/features/2000/600_phone.html*

Ormel, J., Oldehinkel, A. J., & Brilman, E. I. (2001). The interplay and etiological continuity of neuroticism, difficulties, and life events in the etiology of major and subsyndromal, first and

recurrent depressive episodes in later life. [Electronic Version]. *American Journal of Psychiatry, 158(6)*, 885-891.

Orne, M. T. (1962). On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist, 17(11*), 776-783.

Page, S. (1999). Two studies of gender and reporting differences with the Beck depression inventory. *Current Research in Social Psychology, 4(5*), 146-159.

Page, S. & Bennesch, S. (1993). Gender and reporting differences in measures of depression. *Canadian Journal of Behavioural Science, 25(4*), 579-589.

Page, S. & Yates, E. (1975). Effects of situational role demands on measurements of attitudes about mental illness. *Professional Psychology, 6(2),* 175-181.

Rosenthal, R. (1976). *Experimenter effects in behavioral research*. New York: Halsted Press.

Vredenburg, K., Krames, L., & Flett, G. L. (1986). Sex differences in the clinical expression of depression. *Sex Roles, 14(1-2*), 37-49.

Webb, E., Campbell, D., Schwartz, R., & Sechrest, L. (1966). *Unobtrusive measures*. Chicago: Rand McNally.

AUTHOR NOTES

The present study is based on an Honours B.A. Thesis, submitted by the first author, in the Department of Psychology, University of Windsor, 2006.

APPENDIX

Pearson *r* Correlations of Total Beck, Total Hopkins, and Total Risk (health risk variables summed) by Gender (N = 143)

	Condition	Gender	Total Beck	Total Hopkins	Total Risk
Condition	1	-.253 **	.126	.122	.035
Gender	.253 **	1	-.120	-.045	-.016
Total Beck	.126	-.120	1	.218 *	.257 **
Total Hopkins	.122	-.045	.218 *	1	.379 **
Total Risk	.035	-0.16	.257 **	.379 **	1

* Significant at $p < .05$

** Significant at $p < .01$

AUTHOR BIOGRAPHY

Sarah Jane Renaud completed her honors B.A. in the Department of Psychology, University of Windsor, in 2006. She is now a doctoral student in the Department of Educational and Counselling Psychology, McGill University.

Stewart Page, Ph.D., received his Ph.D. from the University of Toronto, and is University Professor of Psychology, University of Windsor. His interests are primarily in social and educational psychology. Contact him at page@uwindsor.ca.