

# CURRENT RESEARCH IN SOCIAL PSYCHOLOGY

---

Volume 4, Number 5  
Submitted: April 21, 1999  
Resubmitted: May 25, 1999  
Accepted: June 3, 1999  
Publication date: June 17, 1999

## TWO STUDIES OF GENDER AND REPORTING DIFFERENCES WITH THE BECK DEPRESSION INVENTORY

Stewart Page  
University of Windsor

[146]

[147]

### ABSTRACT

*In Study 1, 175 participants, 110 females and 65 males, responded to items from the Beck Depression Inventory. Some (69) completed items under testing conditions portrayed as measuring depression. Others (106) completed the same items in a condition portrayed only as measuring reactions to the "hassles" of daily life. In Study 2, 177 participants, 109 females and 68 males, responded to the Beck items under explicit cues for measurement of depression. The notion that participants would respond differently according to gender role congruence of the testing situation was explored in both studies, for example, that males would endorse test items in a less "depressed" direction when items were presented explicitly as measuring depression, but would endorse more depressive content when items were presented in a context of daily hassles. In Study 1, the main effects for testing cues and for participant gender were nonsignificant, but a significant interaction effect showed that although there were no such differences for males, the mean score for females was higher in the depression than in the hassles condition. In Study 2, the mean score for females was significantly higher than that for males. Results seemed to give further support for the idea that reporting differences are relevant to interpretation of depression and possibly other types of personality assessment, that is, that the sexes respond differently depending on the explicitness and role relevance of surrounding test cues.*

[147]

[148]

### INTRODUCTION

The interpretation of gender differences in depression has become the focus of much recent discussion (e.g., Nolen-Hoeksema, 1987; Page, 1990; Stoppard 1989). At present, the nature and

extent of gender differences in depression stands as a contentious, complex, and largely unresolved matter. Stoppard (1989), for example, maintained that women are clearly more depressed than men and also that this difference was authentic and not attributable to differences in how the genders might respond to typical measures of depression. This view has been challenged, for example, by Pyke and Toukmanian (1989, p. 63), who described the gender difference in depression as, among other things, a "fuzzy generalization." Rothblum (1983) maintained, further, that the typically higher rates of treatment for depression in women were attributable mainly to the presence of lower depression specifically among married men. Nevertheless, at present, the reality and interpretation of gender differences in depression remain matters for future research and additional debate.

One aspect of this debate, but one which has received surprisingly little research, is the question of reporting differences. That is, males might be as depressed generally as females, but less likely to portray themselves in this light on commonly used measures. This possibility has not always been taken seriously, and seemingly has been discounted by some writers on depression. Yet others have pursued it more vigorously. Meichenbaum, Price, Phares, McCormick, & Hyde (1989), for example, stated that "Women may not be more depressed than men" (p. 277) because men may not reliably report, or through various ways may mask, their depression (see also Hammen, 1989).

Freudenberger (1987), and Kleinke, Staneski, & Mason (1982) have also discussed strategies by which some males might attempt to camouflage or sublimate depressive symptoms. Vredenburg, Krames, & Flett (1986) found that depressed, male psychiatric patients, more than females, generally endorsed symptoms on the Beck Depression Inventory (BDI) when these were more consistent with the male gender role (see also Goldberg, 1976; Hammen & Padesky, 1977). There is also some evidence that males may conceal the presence of depression out of concern for social rejection (Vredenburg et al., 1986) and because the prevailing image of the depressed man may be still more negative or atypical than that of the depressed woman (Broverman, Vogel, Broverman, Clarkson, & Rosenkrantz, 1970; Hammen & Peters, 1978; Page, 1985; Rosenfield, 1982; Waisberg & Page, 1988).

[148]

[149]

Other literature dealing more explicitly with the social psychology of psychological testing also seems relevant to the issue of reporting differences. It is known that enactment of social roles can affect responses in a testing situation (Kroger & Turnbull, 1975). That is, participants may be induced to respond to test items in terms of their beliefs about the purposes of testing. They may tend, for example, to reject items whose content is perceived as incongruent with self conceptions or with appropriate gender role enactment (e.g., Page, 1990; Rosenthal & Rosnow, 1984; Vredenburg et al., 1986; Waisberg & Page, 1988).

Although the possibility of gender-based reporting differences has thus been raised in the relevant literature on depression, still relatively little focused or systematic research has been conducted on the issue. Unfortunately, therefore, we do not presently have a great deal of research-based information about conditions under which these differences might occur. In the present context, the hypothesis was explored that the measurement of depression in males might

be affected by their non-endorsement of items explicitly described and labeled as measuring depression. Page & Bennesch (1991), for example, found that while the BDI scores of females were unaffected by changes in role demands of the testing situation, males showed significantly lower scores where they believed depression was being measured, compared to when they believed the items were measuring their views about daily hassles, without mention of depression. To the extent that males respond in this fashion, the interpretation of studies using depression measures, such as the Beck Depression Inventory, could be affected in that the extent of depression in these individuals will be inaccurately assessed, and, in effect, under-reported. Alternatively, it might be that females may be more sensitive toward or more accepting toward explicit cues for depression in testing situations, and thus show elevated scores in these conditions, relative to scores of males.

To explore these ideas, male and female participants, in Study 1, were thus tested in a condition identified explicitly as involving the measurement of depression (Depression condition). The remainder were tested in a Hassles condition, not identified with the issue or specific measurement of depression.

[149]

[150]

## **EXPERIMENT ONE**

### **METHOD**

#### **Participants and Materials**

Data were drawn from undergraduate psychology classes. Participants were 175, 65 male and 110 female, undergraduate students at the University of Windsor, with mean ages of 21.10 years (females) and 23.71 (males). Most were enrolled in programs within the Faculty of Social Science. Course credit (bonus marks) were given for participation; 2% were in first year, 54 % were in second year, 31% in third year, and 13% were fourth year students.

All participants signed a written consent form, which was detached and kept separate from each participant's questionnaire data. Also on the consent form was a section in which participants could request a full report of the study when it became available. The consent form indicated that participation was voluntary, that participants could decline any items to which they did not wish to respond, and also that they could withdraw from participation at any time. Upon completion of data gathering, participants were given a brief verbal summary of the study (concerning the issue of gender-based reporting differences in measures of depression). Any queries were answered briefly at this point, and participants were reminded that a full report could be obtained later upon request.

All participants completed, during regular class time, a questionnaire composed of the revised 21-item Beck Depression Inventory (BDI) (Beck, Rush, Shaw, & Emery, 1979). BDI total scores (maximum score = 63) were derived as described by Beck and by Derogatis et al. (see also Johnston, 1986; Johnston & Page, 1991).

Although some inconsistencies have been observed in recent studies concerning its factor structure (Lips and Ng, 1986, have reported that a negative view of self factor is the most frequently extracted factor in studies of the BDI with nonclinical populations), the BDI has shown impressive psychometric properties with a variety of normal and psychiatric populations over three decades (Beck & Steer, 1984; Beck, Steer, & Garbin, 1988; Lips & Ng, 1986). Cronbach's alpha statistic was .832 in the present research (study 1) and .824 (study 2). In general, the literature clearly supports the research use of the BDI with non-clinical populations. The BDI was thus selected because of its suitability for present purposes and also because it represents a variety of depressive symptoms.

## PROCEDURE

Manipulation of testing conditions involved both oral and written components. In the Depression condition (N = 69), the questionnaire's face sheet was labeled "Depression Inventory" at the top, and the words "Depression Project" appeared at the top of each succeeding page. No formal script for the verbal instructions was used. Participants were informed that (a) their cooperation was being solicited for a research study concerning the widespread problem of depression, and that (b) the present study was being administered as part of an ongoing depression research program being undertaken by the author, in the department of psychology.

In the Hassles condition (N = 106), no mention whatever was made of depression. Questionnaire items were instead described as part of an ongoing program of research being undertaken by the author, concerning students' views about the hassles of living. The questionnaire's title described it as a measure of the "daily hassles of living, commonly experienced by all people" and succeeding pages contained the heading "Hassles study." Written instructions for completion of items were as used in the Depression condition, except that the standard request, from the BDI, to describe items in terms of "the way you have been feeling..." was replaced with the phrase "how things have been going..."

## RESULTS

Means and standard deviations of scores by condition (possible score range = 0-63) are shown in Table 1. A 2 (Subject Gender) x 2 (Condition) analysis of variance (ANOVA) of BDI total scores, using the SAS GLM (General Linear Models) procedure, with Horst and Edwards' (1982) correction for unequal cell frequencies, showed both main effects to be nonsignificant, that is,  $p < .05$ . The interaction effect, after removal of variance due to main effects, was significant,  $F(1, 171) = 3.77, p < .05$ . Although the other simple effects tests were nonsignificant, a one-way ANOVA, using the GLM procedure, showed that for females, but not for males, BDI scores were significantly higher in the Depression condition ( $M = 13.28$ ) than in the Hassles condition ( $M = 11.56$ ),  $F(1, 108) = 4.17, p < .04$ .

**Table 1. Means and Standard Deviations of Items by Gender, In Both Conditions (Study 1)\***

<b>Cues</b>	<b>Gender</b>	<b>N</b>	<b>M</b>	<b>SD</b>
Depress	Male	19	11.36	3.19
Depress	Female	50	13.28	5.24
Hassles	Male	46	12.23	3.54
Hassles	Female	60	11.56	3.38

\* Higher means indicate greater tendency toward depression.

## **DISCUSSION**

Mean BDI scores for the present sample were comparable though somewhat higher in magnitude than those found with comparable (Canadian) samples by Johnston (1986) (see also Page & Bennesch, 1991; Johnston & Page, 1991). Scores were also slightly higher than found, for example, by Gould (1982) and by Bourque and Beaudette (1982) with university undergraduate samples, and were generally just above the range often considered as approaching that of mild depression (i.e., mean BDI = 10-11)--though substantially lower than typical scores obtained in research with various clinical or outpatient samples (e. g., Beck & Steer, 1984; Vredenburg et al., 1986).

[152]

[153]

The data (see Table 1) showed no overall gender differences in BDI scores, a finding common to several (though not all) studies with university undergraduates cited, for example, in Stoppard's (1989) extensive literature review of depression research literature. Unlike the earlier cited results of Page & Bennesch (1991), in which males reported more depressive content when items were believed only to be measuring hassles, etc., the present sample of males was not significantly affected by differences in the context or ostensible purpose of testing, although their scores were generally lower in the Depression condition. Females, however, did more frequently endorse depressive content when items were described explicitly as assessing depression, that is, it was the females in the present study, and not the males, whose scores were more strongly affected by changes in the testing situation.

In Study 2, it was decided to attempt to replicate and examine further the results from Study 1 in a different fashion, that is, by using the opportunity to test available participants in a large first-year class, and by focusing interest specifically on the issue of reporting differences in a condition involving explicit cues for measurement of depression. The BDI items were thus administered in a Depression condition, similar to that in Study 1, in which participants completed items in a situation portrayed explicitly as measuring depression.

## **EXPERIMENT TWO**

## METHOD

### Participants and Procedure

Participants were 177 first-year undergraduates, 109 females (mean age = 20.25 yrs.) and 68 males (mean age = 21.74 yrs.) enrolled in introductory psychology at the University of Windsor. Administration was completed in regular class time, during the winter semester, during which the second half of the university's "large section" introductory psychology course was in progress. Data were gathered at a point just after the topic of depression had been introduced, that is, in the concluding portion of the class's previous lecture, although no mention had yet been made concerning possible gender differences.

[153]

[154]

Participants responded to the BDI items, which bore the label "Depression Inventory" on the face sheet, and the description "Depression Measurement Project" at the top of each page. Participants were also informed, as in Study 1, that the data would constitute part of an ongoing research program concerning depression among university populations. Other details of test administration were as in Study 1, including the information given to participants, provision of a brief verbal summary, opportunity to obtain a full report at a later time, and procedures regarding consent.

## RESULTS

An analysis of variance (ANOVA), using the SAS GLM procedure, showed that the mean inventory score for females ( $N = 109$ ,  $M = 14.55$ ,  $SD = 4.48$ ) was significantly higher than for males ( $N = 68$ ,  $M = 12.35$ ,  $SD = 7.71$ ),  $F(1, 175) = 5.27$ ,  $p < .02$ .

## GENERAL DISCUSSION

The present results seem to support the view that testing conditions can cause respondents to alter responses to test items according to whether endorsement may be perceived as gender role appropriate or inappropriate. Several studies of depression in male and female college students and other non-clinical samples (see Lips & Ng, 1986; Stoppard, 1989) have also been undertaken in which overall gender differences have generally not been found. However, in much of this research, reporting differences, especially in response to variations in testing situations, have rarely been formally analyzed or considered. Although overall gender differences in personality variables such as depression may not always be found as main effects, it is obviously possible that gender may still interact with some aspect of the testing situation, such as the manipulation employed in the present study, as well as the similar one, cited earlier, used by Page & Bennesch (1991). These effects seem relevant to both genders. For example, aside from depression, it is conceivable that situational variables which precipitate reporting differences may affect responses of females to measures dealing with stereotypically masculine domains such as aggression, antisocial disorders, or substance abuse. In this wider context, Page (1985) and Rosenfield (1982) found that mental health professionals perceived disorders involving substance abuse or aggression as more serious, and more likely to warrant psychiatric

hospitalization, when observed in females. They also perceived disorders involving depression or anxiety states as more serious and more likely to result in hospitalization when observed in males (see also Waisberg & Page, 1988).

[154]

[155]

Situational cues in testing situations vary in explicitness regarding the ostensible purpose or gender role relevance of testing. While such cues in the present research were deliberately made explicit in order to test more definitively for their effects, others may be more subtle or implicit, as in testing situations which typically ask participants to report on "the way you feel today, right now" or to reveal "problems you may have had in the last few months," and so on. In many testing situations, the surrounding cues are unclear, inconsistent, or ambiguous. Depression inventories, for example, may or may not include a title including the word "depression," or be described verbally as involving depression. In some cases, respondents may be given much information but in other cases may be told little or nothing about the background or purpose of testing, etc., until after data are collected. It is possible that variability in these types of factors may contribute toward the lack of consistent gender differences often obtained with depression inventories in previous research. An hypothesis for future research would be that the tendency to respond to test items in terms of perceived social or gender role demands will vary directly with cue explicitness, that is, within a realistic range of explicitness. It is noted that many measures, even if not explicitly labeled as measures of depression, may nevertheless be (or be perceived as) transparent in this regard by respondents, and may therefore be construed as measures of depression nonetheless. The present results show that, to the extent that situational cues convey the notion of depression, responses may be affected accordingly; that is, role demands in testing situations bearing explicit cues for depression may be ones which tend to maximize endorsement of items for females, but minimize it for males.

Again, the literature on depression has concentrated relatively little on these aspects of testing. In line with the present results, however, Stanton, Burker, & Kershaw (1991) found that participants of both genders generally gave lower depression test scores when they were led, in consent forms, to expect they would be followed up for further clinical investigation. Holowaty (1993) found that both male and female participants who were feminine-typed on Bem's Sex Role Inventory scored significantly higher on the BDI, compared to those who were masculine-typed or androgynous. Holowaty's results indicate that participants' responses may thus be affected by role demands related to one's personal gender orientation, as well as from the immediate influence of a particular test-taking situation. Also, Kornblith, Greenwald, Michelson, & Kazdin (1984), interpreting their results in terms of the notion of demand characteristics, found that participants gave significantly lower BDI scores when told that scores measured the severity of clinical depression, compared to when they were told the scores measured common, everyday problems. Unfortunately, however, these authors did not report nor analyze for the factor of participant gender.

[155]

[156]

Certain limitations of the present research must be acknowledged. The present participants were relatively young. Moreover, Gotlib (1984), for example, found that BDI scores were correlated with other measures of maladjustment in student populations, and thus may not be measuring depression alone. It therefore seems important that future research address further how various aspects of testing situations may affect responses not only to measures of depression, but to measures of other forms of maladjustment as well. Also, the present verbal instructions were brief, in that the use of a formal or lengthy "script" was felt to be contrived and unnecessary. It is thus important for future research to examine a given hypothesis with a variety of methods and conditions. This is meant to emphasize the issue of replicating an idea or hypothesis, rather than being limited to exact replications of specific studies (Rosenthal and Rosnow (1984). In any case, the present data seem generally to show that the notion of reporting differences, and its relationship to the concept of gender role, should not be dismissed as easily as some have implied.

### **AUTHOR'S NOTE**

The present study was supported in part by a President's SSHRC Fund grant to the author, from the University of Windsor. Additional information or details of procedures used are available at any time, upon request. I thank Jennifer Scobie and Rosie Page for assistance with the present study. I also thank two anonymous reviewers of *Current Research in Social Psychology* for comments which were helpful in revising this paper.

### **REFERENCES**

- Beck, A., Rush, A., Shaw, B., and Emery, G. (1979). *Cognitive therapy of depression*. New York: Guilford Press.
- Beck, A., & Steer, R. (1984). "Internal consistencies of the original and revised Beck Depression Inventory." *Journal of Clinical Psychology*, 40:1,365-1,368.
- Beck, A., Steer, R., & Garbin, M. (1988). "Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation." *Clinical Psychology Review*, 8:77-100.
- Bourke, P., & Beaudette, D. (1982). "Etude psychometrique du questionnaire de depression de Beck aupres d'un echantillon d'etudiants universitaires francophones." *Canadian Journal of Behavioural Science*, 14:211-218.

[156]

[157]

Broverman, I., Vogel, S., Broverman, D., Clarkson, F., & Rosenkrantz, P. (1970). "Sex role stereotypes and clinical judgements of mental health." *Journal of Consulting and Clinical Psychology*, 34:1-7.

Freudenberger, H. (1987). "Today's troubled men." *Psychology Today*, 21(12):46-48.



Goldberg, H. (1976). *The hazards of being male*. New York: Signet Books.

Gotlib, I. (1984). "Depression and psychopathology in university students." *Journal of Abnormal Psychology*, 93:19-30.

Gould, J. (1982). "A psychometric investigation of the standard and short form Beck Depression Inventory." *Psychological Reports*, 51:1,167-1,170.

Hammen, C. (1989). "Understanding depression in women: The cognitive approach as an ally." *Canadian Psychology*, 30:52-55.

Hammen, C., & Padesky, C. (1977). "Sex differences in the expression of depressive responses on the Beck Depression Inventory." *Journal of Abnormal Psychology*, 86:609-614.

Hammen, C., & Peters, S. (1978). "Interpersonal consequences of depression: Responses to men and women enacting a depressed role." *Journal of Abnormal Psychology*, 87:322-332.

Holowaty, K. (1993). *Self report of depressive symptoms and daily hassles: Effects of Instructional Set, Gender, and Gender Role Orientation*. Unpublished master's thesis, University of Windsor.

Horst, P., & Edwards, A. (1982). "Analysis of nonorthogonal designs: The k factorial experiment." *Psychological Bulletin*, 91:190-192.

[157]

[158]

Johnston, M. (1986). *Life stress theory and role stereotypes as predictors of psychological adjustment*. Unpublished doctoral dissertation, University of Windsor.

Johnston, M., & Page, S. (1991). "Subject age and gender as predictors of life stress, attributional style, and personal adjustment." *Canadian Journal of Behavioural Science*, 23:475-478.

Kleinke, C., Staneski, A., & Mason, J. (1982). "Sex differences in coping with depression." *Sex Roles*, 8:887-889.

Kornblith, S., Greenwald, K., Michelson, L., & Kazdin, A. (1984). "Measuring the effects of demand characteristics on the Beck Depression Inventory Responses of College Students." *Journal of Behavioral Assessment*, 6:45-49.

Kroger, R., & Turnbull, W. (1975). "Invalidity of validity scales: The case of the MMPI." *Journal of Consulting and Clinical Psychology*, 43:48-55.

Lips, H., & Ng, M. (1986). "Use of the Beck Depression Inventory with three non-clinical samples." *Canadian Journal of Behavioural Science*, 18:62-74.

Meichenbaum, D., Price, R., Phares, J., McCormick, N., & Hyde, J. (1989). *Exploring choices*. Glenview, IL: Scott, Foresman.

Nolen-Hoeksema, S. (1987). "Sex differences in unipolar depression: Evidence and theory." *Psychological Bulletin*, 101:259-282.

Page, S. (1985) "Commitment decisions in Ontario and similarity of symptomatology to gender role stereotypes." *Ontario Psychologist*, 17:4-8.

Page, S. (1990). "Depression in men, depression in women, and the depressing nature of theories of depression." *Canadian Psychology*, 31:292-294.

[158]

[159]

Page, S., & Bennesch, S. (1991) "Gender and reporting differences in measures of depression." *Canadian Journal of Behavioural Science*, 25:579-589.

Pyke, S., & Toukmanian, S. (1989). "A fuzzy generalization." *Canadian Psychology*, 30:63-64.

Rosenfield, S. (1982). "Sex roles and societal reactions to mental illness: Labeling of 'deviant' deviance." *Journal of Health and Social Behavior*, 23:18-24.

Rosenthal, R., & Rosnow, R. (1984). *Essentials of behavioral research*. New York: McGraw Hill.

Rothblum, E. (1983). "Sex role stereotypes and depression in women." In V. Franks & E. Rothblum (Eds.), *The stereotyping of women* (pp. 83-111). New York: Springer.

Stanton, A., Burker, E., & Kershaw, D. (1991). "Effects of researcher follow up of distressed subjects: Trade-off between validity and ethical responsibility?" *Ethics and Behavior*, 1:105-112.

Stoppard, J. (1989). "An evaluation of cognitive/behavioural theories for understanding depression in women." *Canadian Psychology*, 30:39-59.

Vredenburg, K., Krames, L., & Flett, G. (1986). "Sex differences in the clinical expression of depression." *Sex Roles*, 14:37-48.

Waisberg, J., & Page, S. (1988). "Effects of gender role stereotypes on judgements of male and female mental illness." *Women and Health*, 10:44-55.

## **AUTHOR BIOGRAPHY**

The author holds the BA and MA degrees from the University of Western Ontario, and a Ph.D. in social psychology from the University of Toronto. Research interests are mainly in social and community psychology, including the social psychology of the psychological experiment and

test-taking situations. He is currently professor, Department of Psychology, University of Windsor, Windsor, Ont., Canada N9B 3P4. E-mail = [Page@Server.Uwindsor.ca](mailto:Page@Server.Uwindsor.ca).

[159]

---