INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

NOTE TO USERS

The original manuscript received by UMI contains pages with print exceeding margin guidelines. Pages were microfilmed as received.

This reproduction is the best copy available

UMI

Running Head: RANDOM BECK DEPRESSION INVENTORY

An Assessment of the Random Version of

the Beck Depression Inventory II

Beth A. Merkley

M.A. Thesis

Lakehead University

Thunder Bay, Ontario

Submitted in partial fulfilment of the degree of

Master of Arts in Clinical Psychology

Supervisor: Dr. D. Mazmanian

Date: August 30, 1997

©Beth Ann Merkley

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-33416-3

Canada

TABLE OF CONTENTS

LIST OF TABLES

Abstract

The present study compared a random version of the Beck Depression Inventory II (BDI-IIR, based on the work of Dahlstrom, Brooks, & Peterson, 1990) with the standard version in a sample of 139 university students. The study also assessed various psychometric properties of the Random version, including test-retest reliability, internal consistency, discriminant validity, and concurrent validity using the standard BDI-II, the Zung Self-Rating Scale (SRS), and the Center for Epidemiological Studies-Depression Scale (CES-D) over a two week period. In addition, the effects of prior exposure to a depression inventory were examined. The BDI-II and its random counterpart did not evidence significantly different total scores, and both were found to have adequate and similar test-retest reliability, internal consistency and validity. Test-retest reliability was .86 for the BDI-II and .75 for the BDI-IIR. Internal consistency was found to be .87 and .85 for the BDI-II and BDI-IIR respectively. Both versions of the BDI-II were found to possess adequate validity with the CES-D and SRS. However, both scales also correlated very highly with the Edwards Social Desirability Scale. Lastly, pre-exposure to depression measures did not result in decreased scores at a later administration.

iv

An Assessment of the Random Version of

the Beck Depression Inventory II

Major Depressive Disorder is a mood disorder that has been recognized for centuries. It affects individuals across culture, gender, age and ethnicity, and has been estimated to account for approximately 10% of all advice-seeking visits to family physicians (Bech, 1992). The lifetime risk of this disorder is approximately 5-12% for men and 10-25% for women (American Psychiatric Association, 1994).

In light of the extensive occurrence of depression, the psychometric properties of the various depression inventories are important. These inventories have numerous purposes, such as the selection of research participants, diagnosis in clinical settings, and monitoring treatment efficacy. As is the case for any psychological test, it is important that depression inventories possess adequate temporal stability or test-retest reliability. Thus, a depression inventory should yield consistent scores when the same individual completes the measure on different occasions (Anastasi, 1988).

A second necessary psychometric property is that of validity, the extent to which a test measures what it was designed to measure-in this case depression. One pertinent aspect of validity is known as concurrent validity, the extent to which a scale correlates with other measures of

the same construct (Gregory, 1996). Although the concurrent validity of various depression measures has been examined frequently (Joseph, Lewis, & Olsen, 1996; Kerner & Jacobs, 1983; Radloff, 1977; Tanaka & Huba, 1987; Tanaka-Matsumi & Kameoka, 1986), few studies have investigated this type of validity using multiple concurrent measures at multiple points in time.

Convergent and discriminant validity are additional aspects of construct validity. A test is said to possess convergent validity when it correlates highly with measures of similar constructs, and to possess discriminant validity when it does not correlate with theoretically unrelated measures (Cambell & Fisk, 1959; Gregory, 1996). While most depression inventories appear to possess adequate convergent validity (Beck, Steer, & Garbin, 1988; Tanaka-Matsumi & Kameoka, 1986), the discriminant validity of many instruments remains open to question. Of particular concern is the fact that many measures of depression are negatively correlated with measures of social desirability (Beck et al., 1988; Cole, 1988; Tanaka-Matsumi & Kameoka, 1986). The socially desirable response style reflects the tendency to ascribe to oneself socially desirable attributes and to present oneself in a favourable light (Ones, Viswesvaran, & Reiss, 1996). In the case of depressed individuals, self-presentation is undesirable and unfavourable. In other words, a respondent might endorse items on the basis of

their degree of social desirability or undesirability rather than the degree to which the item content is true of them (Jackson, 1989).

At present, a large assortment of self-report measures exist to aid in the assessment of depression. These measures include the Beck Depression Inventory (BDI; Beck & Steer, 1987), the recently published Beck Depression Inventory II (BDI-II; Beck, Steer, & Brown, 1996), the Depression Scale of the Minnesota Multiphasic Personality Inventory-2 (MMPI-2D; Butcher, Dahlstrom, Graham, Tellegen, & Kaemmer, 1989), the Hamilton Rating Scale for Depression (HRSD; Hamilton, 1960), the Center for Epidemiological Studies-Depression Scale (CES-D; Radloff, 1977) and the Zung Self-Rating Scale (SRS; Zung, 1965). Despite the abundance of measures, the BDI remains the most utilized depression instrument worldwide (Beck & Steer, 1987; Gotlib, Lewinsohn, & Seeley, 1995; Margo, Dewan, Fisher, & Greenberg, 1992) due to its numerous advantages. These include its validity, relative ease of administration, the short duration of time required to complete the survey, its ability to yield a quantified evaluation of the level of depression, and the ability to obtain information which may have been overlooked in a clinical interview (Schnurr, Hoaken, & Jarrett, 1976).

Test-Retest Reliability

Numerous studies have sought to determine the test-retest reliability of the BDI over varying lengths of time

and across diverse populations. Test-retest reliability coefficients range from .60 in college students after a one week period (Hatzenbuehler, Parpal, & Matthews, 1983) to .90 in psychiatric and non-psychiatric elderly participants over a range of 6 to 21 days (Gallagher, Nies, & Thompson, 1982). In terms of the student sample, same-day coefficients have been placed at .83 and .81 (Hatzenbuehler et al., 1983). Longer-term coefficients found for nondepressed college samples were .78 after one week (Oliver & Burkham, 1979), .78 over a three week interlude (Oliver & Burkham, 1979), .75 after one month (Pehm, 1976), and .74 ensuing a three month period (Miller & Seligman, 1973). The BDI also has good test-retest reliability in agoraphobic patients (.80, Michelson & Mavissakalian, 1983), and hospitalized adolescents (.74, Strober, Green, & Carlson, 1981).

Some concerns have been raised regarding the test-retest reliability of the BDI, however (Boyle, 1985; Hatzenbuehler et al., 1983). One difficulty is related to the normal fluctuations inherent in the course of depression, whereby a change in severity may indicate a modification in depressive symptomatology, rather than poor test-retest reliability. Thus, scores may be modified as a result of the normal fluctuations inherent in the course of depression. Moreover, discrepancies in the test-retest reliability of depression inventories have been readily observed between psychiatric and non-psychiatric samples

whereby the former manifests greater lability, lending credence to the speculation that natural fluctuations in the course of depression serve to decrease reliability estimates in clinically depressed samples (Beck & Steer, 1987; Beck et al. 1988; Gallagher et al., 1982).

When the aforementioned studies are examined closely, it appears that as the elapsed time between administrations increases, test-retest coefficients begin to decrease, although only by relatively small amounts. A longer delay between BDI administrations may result in a more modest test-retest reliability coefficient as participants who recomplete the BDI after a short interlude may be more likely to recall their prior responses and to subsequently alter their current answers. Hammen, Mayol, deMayo, and Marks (1986), however, postulate that depressive symptomatology will be more labile across longer interludes than across a single day due to the unstable nature of depression. Nevertheless, the evidenced decrease in test-retest coefficients as a function of retest interval is not specific to measures of depression and is a common event with all repeatedly administered measures regardless of content (Graziano & Raulin, 1993).

The Effects of Prior Exposure to the BDI

A further concern is that repeated testing with the BDI appears to result in reduced scores on subsequent BDI administrations. For example, Beck et al. (1988) and

Hatzenbuehler et al. (1983) found that retest scores were significantly lower than scores obtained at the initial testing. In regards to the latter study, these results were found for participants retested either during the same day or one week later, and only among those who initially scored in the mild and moderate ranges. In addition, Hammen (1980) found an average drop from 18.37 to 10.29 points when college students were retested within 2-3 weeks of the initial administration, again demonstrating a reduction in severity scores on a subsequent administration of the BDI.

In a similar manner, Atkeson, Calhoun, Resick, and Ellis (1982) repeatedly assessed rape victims using the BDI and the HRSD (Hamilton, 1960) at several points following the assaults. Control groups comprised of similar rape victims were also assessed but at only one point in time. Upon comparison of the repeatedly assessed victims with the control group, a significant difference was found at 4 months post-rape, and a near significant difference was found at 2 and 8 months post-rape. Specifically, the repeatedly assessed group exhibited diminished BDI scores at the three comparison points, suggesting that repeated assessments with the BDI may lead to a decrease in severity scores. Interestingly, however, no difference between groups was found with the HRSD.

Further information on the effects of prior test exposure was provided by Lambert, Hatch, Kingston, and

Edwards (1986), who conducted a meta-analytic study involving the BDI, the SRS and the HRSD on individuals receiving treatment for depression. Although the findings were confounded by discrepancies across studies in assessment interludes and forms of treatment, it was observed that when either the BDI or the SRS were administered on at least one occasion beyond pre-and posttreatment sessions, the pre-to posttreatment effect size increased somewhat for those who received more than two administrations when compared to those who received only pre-and post treatment assessment. Interestingly, although the effect size of the HRSD was larger, it remained constant over time. Nevertheless, despite the lack of obvious differences in treatments, it appeared that exposure to depression measures enhanced treatment effects over and above that expected from treatment alone. Thus, prior exposure was surmised to result in decreased scores at later administrations beyond the effects of any treatment procedures.

Perhaps one of the most relevant studies pertaining to the effects of repeated administrations is that by Hatzenbuehler et al. (1983) in which college students were evaluated on two separate occasions. Participants were administered either the BDI or the SRS at Time 1, followed by the BDI later the same day. Participants who completed the BDI on both occasions exhibited significantly lower

scores at the second administration when compared to those who completed the SRS during the initial testing session, thus indicating that completing the same test more than once may result in lowered scores at later testing sessions.

When Hatzenbuehler et al. (1983) readministered the BDI to a new sample either a few hours or seven days ensuing the first administration, all participants evidenced significantly lower scores at the time of retest, regardless of the retest interval. The authors proposed that, given the short re-test interval, fluctuations in scores may be due to participants recalling their prior responses resulting in subsequent modification of their later answers. This being the case, it would be beneficial to determine if a similar decrease in scores will be evident if the interval between testing times is increased to a point where the participants will no longer be able to recall their prior responses.

Hatzenbuehler et al. (1983) have also speculated that the drop in scores evidenced at the second administration may be the result of a change in expectancies and/or due to increased knowledge of depressive symptomatology garnered from prior exposure to the BDI. Although participants may become more cognizant regarding the symptoms of depression, it cannot be stated with certainty that the increase in knowledge will result in a wilful modification of future responses. Furthermore, it cannot be properly ascertained that participants recall actual items from the scale, nor

that they universally attempt to change their responses. Hence, the aforementioned statement by Hatzenbuehler et al. (1983) remains speculative at best.

In a related study, Mark, Sinclair, and Wellens (1991) found that completing the BDI results in an alteration of reported mood. Specifically, upon completion of the BDI, depressed participants expressed an increase in negative affect when assessed with a 20-item mood measure (presumably constructed by the authors), whereas non-depressed participants reported an elevation in positive feelings. This was postulated to be due to the comparison of current and past moods, resulting in a subsequent priming of negative affective schemas in depressed participants. If depressive schemas are indeed primed in previously depressed individuals, then these participants would be more likely to endorse options indicative of depression when retested after a succinct interval of time, providing that the interval is short. The accuracy of this speculation can be ascertained by comparing participants whose depression scores fall in the upper or lower third of the depression continuum. That is, individuals scoring in the upper extreme of the distribution would be expected to evidence an increase in depressive severity, whereas those in the lower extreme would exhibit either no change or increased depression scores (regression to the mean) at later administrations.

Social Desirability and its Relation to Depression

Inventories

An additional factor which may affect the validity of self-report measures of depression is social desirability, the tendency to present oneself in an overly positive and socially desirable manner (Murphy, 1993). Specifically, the use of response sets can limit the accuracy of the scores obtained with the BDI-II and its predecessors, as participants may not admit to and subsequently not endorse items which they deem to reflect socially undesirable behaviours, attitudes, or thoughts (Murray, 1973). Conversely, individuals who are depressed may rather endorse items on the basis of undesirability and thus ascribe undesirable characteristics to themselves. Despite the potential negative effects that social desirability may have on the BDI and BDI-II, it is often left unevaluated. Beck (1972) reported a significant negative relationship (-.55) between BDI scores and social desirability. Similar results were found by Cole (1988) with respect to social desirability and depression as measured with the SRS and the MMPI-D (Hathaway & McKinley, 1942). In addition, social desirability has been found to correlate negatively with related constructs such as hopelessness (Linehan & Nielson, 1981, 1983; Mendonca, Holden, Mazmanian, & Dolan, 1983; Petrie & Chamberlain, 1983), and suicide ideation (Holden, Mendonca, & Mazmanian, 1985).

The manner in which these results apply to college

populations is of particular interest. Due to the relative
ease of obtaining participants, college students are often
employed in research studies (Hatzenbuehler et al., 1983;
Tanaka & Huba, 1987). However, it is also at this
developmental stage that social desirability plays an
important role in everyday events, especially in the form of
egocentrism (Dahlstrom et al., 1990; Santrock, 1993).
Hence, to the extent that students may be responding to the
desirability of items rather than its content per se,
depression scores in this sample may be attenuated and may
therefore not reflect the true severity of depression.

A different relationship is evidenced among depressed
individuals. Contrary to students, depressed individuals
tend to portray less social desirability and present
themselves in a more negative manner, resulting in higher
depression scores (Abramson, Seligman, & Teasdale, 1978;
Beck, 1976). The nature of the relationship between
depression and social (un)desirability, particularly the
question of whether social desirability represents error or
is rather an intrinsic and legitimate aspect of depression,
is still debated (Katz, Shaw, Vallis, & Kaiser, 1995).

When Tanaka-Matsumi and Kameoka (1986) compared the SRS
and the BDI with both the Marlowe-Crowne Social Desirability
Scale (MCSDS; Crowne & Marlowe, 1960) and the Edwards Social
Desirability Scale (ESDS; Edwards, 1970), the relationship
between depression scores and social desirability was found

to vary primarily on the basis of which social desirability scale was employed. Specifically, the ESDS exhibited a significant negative correlation with both the BDI and the SRS (-.54 and -.58 respectively), whereas the relation between the two depression measures and the MCSDS was not significant. It should be noted, however, that the ESDS contains numerous items which reflect depressive symptomatology. In a separate study, Holden and Fekken (1989) further found that the Jackson Social Desirability Scale of the Personality Record Form (PRF) (Jackson, 1989) and the ESDS both defined one social desirability factor, whereas the MCSDS formed a separate factor.

The Random Form of the BDI

Another important concern regarding the BDI-II and its predecessors is their potential susceptibility to perseverative response sets (Beck & Beamesderfer, 1974; Dahlstrom et al., 1990). The BDI-II employs a consistent response order for each of its 21 items whereby the options for each item increase consistently in severity. That is, each response option within a single item grows increasingly severe in nature, ranging from, for example, "not at all" to "all of the time". As the increase in severity is uniform across all 21 items, it is possible to select the most or least pathological option without reading the options or considering them carefully. Hence, it does not require great effort to select a desired severity nor to present a

favourable or malingering response set (Dahlstrom et al., 1990).

In order to test the above hypothesis, Dahlstrom et al. (1990) devised a randomized variant of the 1961 version of the BDI, in which severity did not increase uniformly (BDI-R). Rather, some items began with the most severe option, whereas others commenced with a less severe option. In addition, the order of severity within each of the 21 items was varied, thus ensuring that the response options were not presented in a standardized manner.

Dahlstrom et al. (1990) administered this measure to a sample of college students. The original version of the BDI was also administered, in addition to a backwards version whereby the most pathological options were presented first. The MMPI-D was also employed to discern whether differences in depression existed among the three groups receiving the various forms of the BDI.

Participants who completed the BDI-R exhibited significantly higher scores compared to those who had completed either the original or backwards versions of the BDI. Furthermore, no significant differences were found between the latter two versions of the BDI, suggesting that the backwards version is similar to the original in terms of symptom severity and the use of response sets. In light of the discrepancy in scores between the BDI-R and the original and backwards versions of the BDI, it would appear that the

BDI-R might require participants to peruse the descriptors more thoroughly, thus reducing the use of response sets.

Of further interest is the lack of significant differences among the MMPI-D scores for the three BDI groups. Although the different BDI forms resulted in discrepant levels of depression, all three groups exhibited the same degree of depressive symptomatology when measured with the MMPI-D. Hence, it remains uncertain if the use of the BDI-R results in a more authentic assessment of depression, or rather, if it spuriously raises the obtained level of depressive severity. Despite the potential applicability of these findings to research and clinical practice, several apprehensions remain. For example, the study has yet to be replicated, and the BDI-R has not been examined on male subjects.

In summary, the BDI and BDI-II have been used in numerous populations to assess depressive symptomatology and severity. Despite their wide use, some concerns remain. These include questions about test-retest reliability, possible reduction in scores due to prior exposure, and the possible effects of response sets on obtained depressive severity. It would be beneficial to garner further knowledge regarding these issues.

The Present Study

The primary purpose of the present study was to replicate and expand the findings of Dahlstrom et al. (1990)

and examine the reliability and validity of the recently published BDI-II. Specifically, the present study sought to determine if the random version of the BDI-II (BDI-IIR) would evidence elevated depression scores when compared to scores obtained using the BDI-II. Based on the results of Dahlstrom et al. (1990), it was hypothesized that the BDI-IIR would result in higher depression scores when compared to its counterpart, as the use of response sets would be attenuated (Hypothesis 1).

In addition, few studies have examined the psychometric properties of the BDI-II. Hence, the present study also sought to determine the validity of the BDI-II and the BDI-IIR utilizing the CES-D and the SRS. It was hypothesized that both versions of the BDI-II would possess adequate concurrent validity as evidenced through a positive correlation with the CES-D and SRS (Hypothesis 2). As individual measures contain error, reduction of error variance would be beneficial. This can be accomplished through the use of aggregation. Aggregation is a procedure which combines multiple measurements, resulting in reduced error variance and a more stable and unbiased estimate when compared to single measures (Epstein, 1979; Ossenkopp & Mazmanian, 1985). Aggregation of the two measures (CES-D & SRS) should provide a more stable measure of depression than either scale alone. A further purpose was to ascertain the test-retest reliability of the BDI-II and the BDI-IIR. It

was hypothesized that both the BDI-II and BDI-IIR would evidence adequate test-retest reliability (.70 or greater, Murphy, 1993) (Hypothesis 3).

In accordance with Boyle's (1985) recommendation that stable levels of depressive severity be utilised for the assessment of test-retest reliability, a non-clinical population was employed. This population allowed for a more accurate assessment of test-retest reliability due to their greater stability in affective states. Thus, a university sample permitted a more accurate assessment of the test-retest reliability of the BDI-II and BDI-IIR without having to partial out the effects of treatment efficacy or the natural fluctuations inherent in the course of depression. An additional objective was to examine the potential impact of socially desirable response styles on BDI-IIR and BDI-II scores. Both versions were hypothesized to correlate negatively with social desirability (Hypothesis 4).

Lastly, the present study sought to determine if prior exposure to a depression survey would result in a decrease in scores at a later testing. By comparing the scores of participants who previously completed a depression inventory to those who had not, it is possible to determine whether repeated exposure serves to reduce future scores. The use of a control group offers the advantage of examining the effects of pre-exposure in different subjects. It is also possible to determine the effects of exposure by examining

the scores of participants who completed the same measure on two occasions. Specifically, based on previous findings by Hammen (1980), Hatzenbuehler et al. (1983), and Lambert et al. (1986), it was hypothesized that participants who had previously completed a depression survey would exhibit reduced scores when compared to those who had not completed any prior surveys. Similarly, it was hypothesized that participants who completed the same measure twice would evidence a significant reduction in scores from Time 1 to Time 2 (Hypothesis 5).

<div align="center">Method</div>

<u>Participants</u>

Participants consisted of 142 students enroled in Introductory Psychology classes at a small central Canadian university. Participants consisted of 40 males (29.2%) and 97 females (70.8%), with a mean age of 21.13 years (range 18-43 years). Two participants failed to indicate their sex. Two subjects were removed from the study due to previous participation at an earlier date. One participant was removed from the study after the second session as her testing protocol for the first session could not be identified. Hence, participants consisted of 139 students during the first testing session and 92 at the retest session. Participants gained up to two bonus points in the course for engaging in the study and their participation was strictly voluntary.

Instruments

Beck Depression Inventory II (Original Version)

The BDI-II (Appendix A) has an internal consistency of .92 in outpatient samples and .93 in student samples (Beck et al., 1996). The measure consists of 21 items which assess the cognitive, affective and experiential components of depression (Beck et al., 1996; Lambert et al., 1986). Each item consists of four response options whereby the severity of the options increases in an ascending manner. Item responses are summed to obtain a total score, which is then used to classify individuals into one of four depressive categories: a total score of 0-13 indicates the absence of depression, a sum between 14 and 19 is associated with mild depression, 20-28 represents moderate depression, and a score of 29-63 depicts the presence of severe depression. The BDI-II can be completed in approximately 5 to 10 minutes (Beck et al., 1996).

Randomized Beck Depression Inventory II

The BDI-IIR (Appendix B) comprises the same 21 items found in the BDI-II. However, rather than presenting the severity options for each item in a standardized order, the options are presented in a randomized fashion. Items are presented in the same sequence as the BDI-II and the specific randomized order of the options was based on those delineated by Dahlstrom et al. (1990) (Appendix C). However, the multiple responses per item occurring in the

1961 version of the BDI (e.g., 1A, 1B) were eliminated in an alternating manner in order to account for their removal from the recent publication of the BDI-II, with the exception of Items 16 and 18.   Item scores were converted to their original value prior to summation.   Thus, the total score is obtained in the same manner as the BDI-II.   The BDI-R has been found to correlate with the Harris-Lingoes Depression Subscales of the MMPI-D to the same degree as the original BDI (Dahlstrom et al., 1990).

Zung Self-Rating Scale

The SRS (Appendix D) is a widely used measure of depression consisting of 20 items that assess the pervasive affective, physiological and psychological symptoms of depression (Zung, 1965).   Half the items are worded in a positive direction, whereas the other half are worded negatively, thereby reducing the effect of response sets. All response options, "a little of the time, some of the time, a good part of the time, and most of the time", are placed in the same order regardless of the direction of the item wording.   To ensure that the same time frame was employed across all measures, participants were asked to base their assessments on the past two weeks, as is done with the BDI-II.

Item values range in severity from 1-4 and the total score is obtained by summing the scores of all 20 items The total score is then multiplied by 1.25 and divided by

100, resulting in a range of scores from .25-1.00. Scores
below .50 are considered normal, whereas scores between .50-
.59 are indicative of mild-moderate depression, those in the
.60-.69 range denote moderate-severe depression, and scores
.70 or higher are considered representative of severe
depression. The internal consistency of the SRS has been
estimated at .88 for depressed individuals and the split-
half reliability is .94 (Gabrys & Peters, 1985).

Center for Epidemiological Studies Depression Scale

The CES-D (Appendix E) is a 20 item self-report
instrument designed to measure current levels of depressive
symptomatology (Radloff, 1977). Four items are worded in
the positive direction and 16 items are worded in the
negative direction. Scores range from 0 to 60, with higher
scores being indicative of greater symptomatology. The
internal consistency of the CES-D has been estimated to be
.85 in general populations and .90 in psychiatric samples
(Radloff, 1977). The CES-D has also been found to correlate
.83 with the Symptom Checklist-90 Depression scale in
psychiatric samples (Radloff, 1977), and .69 with the BDI in
college samples (Tanaka & Huba, 1987).

Edwards Social Desirability Scale

The ESDS (Appendix F) contains 39 items derived from
various scales of the MMPI which were judged by a sample of
152 students to reflect socially desirable traits (Edwards,
1957). The items assess social desirability using a

True/False format (Tanaka-Matsumi & Kameoka, 1986). A score of one point is given if the participant answers "true" to a socially desirable item or "false" to a socially undesirable item. All points are totalled and higher scores indicate greater social desirability. The corrected split-half reliability has been estimated at .83 (Edwards, 1957), and the ESDS has been found to correlate .71 with the Jackson Social Desirability scale of the PRF (Holden & Fekken, 1989).

Demographic Survey

A demographic survey (Appendix G) created by the author was employed to gather information regarding the participants' age, sex and psychiatric history.

Consumer Survey

The consumer survey (Appendix H) was created by the author to assess product selection for the purpose of engaging the participants in the control groups in an activity, thus reducing their exposure to the scales being completed by their neighbours. The survey also ensured that control participants spent an equal amount of time completing measures in the testing situation as did their neighbours.

Procedure

Participants were approached during class time and given a brief recruitment speech (Appendix I). The students were then asked if they would like to volunteer as a

participant and were subsequently requested to sign up for a research session. The students were also informed that should they decide to volunteer, they would be asked to attend two sessions lasting approximately 25 to 40 minutes each in order to receive their bonus points. One bonus point was given per session attended. Sessions were held Monday to Friday at various intervals between 12:30 and 5:00pm during several week-long periods. One initial and retest session were held following a midweek night class.

All participants received the same standardized instructions (Appendix J) prior to each testing session, and were informed that their participation was strictly voluntary and that they may withdraw from the study at any time. Participants also received a consent form (Appendix K) and instructions for the completion of a package of questionnaires on both occasions, and were again reminded that they may withdraw from the study at any time. The consent form further indicated that research data would be kept for seven years and that a summary of the research findings could be obtained upon completion of the study.

Participants were informed that their responses would be kept confidential and that their student numbers would be utilized solely for the assignment of bonus points. Lastly, participants were informed that their student numbers would be deleted when the data was entered into the data bank and that their answers would therefore be anonymous.

In order to assess the test-retest reliability of the various depression measures, participants were asked to complete a package of questionnaires at the initial session (Time 1), and again two weeks after the initial testing session (Time 2). Upon completion of the questionnaires at Time 1, participants were asked to sign up for the subsequent session and were requested to return on the same day of the week for the Time 2 session, whenever possible. On average, there were 6 to 10 students per testing session.

During the first administration, all participants randomly received one of three packages to complete (Table 1). All packages contained the ESDS and the Demographics Questionnaire. The two depression packages also contained the SRS, CES-D and either the BDI-II (package 1) or BDI-IIR (package 2). The control group received the consumer survey instead of the depression scales (package 3). In addition, each package included a small card which was to be kept by the participant and brought to the second testing session. The cards were colour coded according to the type of package received at Time 1. The informed consent forms returned to the researcher were also colour coded in a similar manner and were subsequently brought to the later testing sessions by the investigator.

At the retest session, participants were asked to complete one of two possible packages of questionnaires. Package 1 contained the ESDS, SRS, CES-D and the BDI-II.

Table 1

Measures Included in each Package[a]

Package   Measures

Time 1

1   Consent Form, Demographics Questionnaire, ESDS,

        SRS, CES-D, BDI-II

2   Consent Form, Demographics Questionnaire, ESDS,

        SRS, CES-D, BDI-IIR

3   Consent Form, Demographics Questionnaire, ESDS,

        Consumer Survey

Time 2

1   ESDS, SRS, CES-D, BDI-II

2   ESDS, SRS, CES-D, BDI-IIR

Note. [a] ESDS represents the Edwards Social Desirability

Scale. The SRS is the Zung Self-Rating Scale. CES-D

signifies the Center for Epidemiological Studies-Depression

Scale. The BDI-II represents the Beck Depression Inventory

II and the BDI-IIR is the Random version of the Beck

Depression Inventory.

Package 2 contained the ESDS, SRS, CES-D and the random
version of the BDI-II (Table 1). The same standardized set
of instructions were given. Identifying information at the
time of retest consisted of only the participants' student
number, professor and class section. The two packages were
distributed according to which of the three packages was
received at Time 1, as determined by the cards retained by
the participants. For example, the first female to present
a Time 1-package 1 card was given package 1 at Time 2,
whereas the second female with a Time 1-package 1 card
received package 2 at Time 2. The same procedure was
performed with the Time 1-package 2 and Time 1-package 3
cards retained by the participants. This ensured an equal
distribution of participants across instruments and further
aided in the balancing of the distribution of males and
females.

Furthermore, some participants completed the same
measure on both occasions whereas others completed a
different version of the BDI-II at each session, thus
creating counterbalanced groups for order of administration.
Due to the removal of the control packages after Time 1,
only the two depression packages were utilized during the
retest session, thus creating two control groups (Control-
BDI-II & Control-BDI-IIR). In addition, several
participants failed to return for the second testing
session, thus creating an additional group. This procedure

resulted in the formation of seven groups in total (Table 2).

Immediately after completing the questionnaires at Time 2, a description of the study (Appendix L) was distributed to all students. A contact number for the researcher and research supervisor were provided, if any questions arose regarding the study. Envelope labels were also made available for participants who desired a research summary following the completion of the statistical analysis. Participants were also revisited during class time and given the opportunity to receive a research summary.

Lastly, in accordance with university policy regarding research procedures, it was arranged that all raw data remain stored in a secured location for a minimum of seven years. Upon completion of my degree, provisions will be made to store the data in a secured environment until the remainder of the seven years has elapsed.

Results

Equivalency of Groups

Prior to analyses, three separate 7x2 chi-square analyses were performed to examine whether confounds existed between the various groups. The first chi-square compared the sex ratio among the seven groups (Table 2). No significant difference was found, $\underline{X}^2(6) = 2.23$, $\underline{p} = .90$. A chi-square was also performed to examine the prevalence of psychiatric disorders among the groups. No significant

Table 2

Group Divisions and Measures Completed

Group	Time 1	Time 2	n
1	Package 1[a]	Package 1	13
2	Package 2	Package 2	19
3	Package 2	Package 1	15
4	Package 1	Package 2	15
5[b]	Package 3	Package 1	16
6[b]	Package 3	Package 2	15
7[c]	Package 1	--	18
	Package 2	--	14
	Package 3	--	14
		Total:	139

Note. [a] Package 1 contained the ESDS, SRS, CES-D and BDI-II.

Package 2 contained the ESDS, SRS, CES-D and BDI-IIR.

Package 3 contained the ESDS and the Consumer Survey.

[b] Participants in the control group at Time 1.

[c] Participants who did not return for the retest session.

differences were found, $X^2(6) = 6.34$, $p = .39$, indicating that the groups were similar in regards to prevalence of psychiatric disorders. The final chi-square examined whether the groups differed in regards to the number of participants with a prior diagnosis of depression. No difference was found, $X^2(6) = 11.77$, $p = .07$ indicating that the groups were similar in regards to prior diagnosis of depression.

To ensure that all groups were equal in depressive severity prior to analyses, one-way ANOVAs were performed on the CES-D and SRS scores, respectively, for the five groups who received depression scales at Time 1 (see Table 2). An alpha level of .01 was employed to reduce Type I error. No significant differences were found for the CES-D, $F(4, 82) = .60$, $p = .66$, or the SRS, $F(4, 89) = .60$, $p = .66$. Furthermore, no differences were found between the seven groups on the ESDS, $F(6, 132) = .21$, $p = .98$, indicating that the groups did not differ in social desirability. The lack of significant differences indicate that random assignment was effectively implemented, that participants were equivalent with respect to the variables of interest and that participants who failed to return for the second session were not significantly different from those who did.

In addition, examination of the univariate distributions (Table 3 and Table 4) indicated that participants scored within three standard deviations above

or below the mean on all measures of depression with only four exceptions. One outlier was found on the BDI-II at Time 2 (BDI-II score = 37, z = 3.09), another on the CES-D at Time 1 (BDI-II score= 49, z = 3.52), and one on the BDI-IIR at Time 1 (BDI-II score = 33, z = 3.53) and Time 2 (BDI-II score = 49, z = 4.12) respectively. Upon closer examination, it was discovered that the outliers on the BDI-II and CES-D were due to the same participant. This participant, as well as the participant representing the outlier on the BDI-IIR at Time 2, had been previously diagnosed with depression.

The issue of retaining or excluding the BDI-IIR Time 2 outlier required careful consideration, especially in view of the relatively small sample sizes. To examine precisely the effect of this outlier, a number of initial analyses were conducted with the outlier excluded. For example, the correlation between the BDI-IIR at Time 1 and Time 2 was .75 with the outlier included and .75 with the outlier excluded. Similarly, the correlation between the BDI-IIR at Time 2 and the SRS at Time 2 was .74 with the outlier included and .77 with the outlier excluded. An ANOVA with the outlier included resulted in $F(1, 32)$ = 2.47, p = .13, whereas the results were $F(1, 31)$ = 1.12, p = .30 when the outlier was excluded. Consequently, it was decided to retain the outlier in all analyses reported.

Psychometric Properties

Table 3

Means and Standard Deviations for the Measures (Time 1)

Group	SRS	CES-D	ESDS	BDI-II	BDI-IIR
1	17.54	16.46	29.08	10.15	--[a]
	(8.04)	(11.59)	(5.24)	(7.88)	--
2	15.47	13.78	29.37	--	8.00
	(8.08)	(8.35)	(5.79)	--	(7.41)
3	16.07	17.13	28.40	--	10.40
	(8.37)	(12.73)	(5.60)	--	(7.65)
4	16.47	14.33	29.13	7.53	--
	(8.66)	(10.77)	(4.79)	(6.60)	--
5	--	--	27.81	--	--
	--	--	(5.50)	--	--
6	--	--	28.47	--	--
	--	--	(7.09)	--	--
7	13.84	13.00	29.37	8.17	7.36
	(7.93)	(7.84)	(5.83)	(5.02)	(5.26)

Note. Standard deviations appear in brackets. SRS is the Zung Self-Rating Scale. CES-D is the Center for Epidemiological Studies-Depression Scale. ESDS is the Edwards Social Desirability Scale. BDI-II is the Beck Depression Inventory II. BDI-IIR is the Beck Depression Inventory II-Random version.

[a] Indicates that the measure was not administered to this group.

Table 4

Means and Standard Deviations for the Measures (Time 2)

Group	SRS	CES-D	ESDS	BDI-II	BDI-IIR
1	18.08	18.00	28.46	11.15	--
	(8.41)	(11.11)	(5.62)	(7.34)	--
2	15.00	12.95	29.16	--	7.47
	(9.21)	(9.37)	(6.35)	--	(7.06)
3	16.13	17.53	28.40	10.00	--
	(8.05)	(12.05)	(6.32)	(9.01)	--
4	14.47	12.73	28.87	--	8.07
	(8.68)	(6.56)	(5.01)	--	(8.53)
5	16.88	19.81	28.25	12.63	--
	(9.56)	(12.20)	(5.74)	(8.72)	--
6	17.53	20.80	27.20	--	12.87
	(8.85)	(14.41)	(7.20)	--	(12.71)
7	--	--	--	--	--

Note. Standard deviations appear in brackets. SRS is the Zung Self-Rating Scale. CES-D is the Center for Epidemiological Studies-Depression Scale. ESDS is the Edwards Social Desirability Scale. BDI-II is the Beck Depression Inventory II. BDI-IIR is the Beck Depression Inventory II-Random version.

To ascertain the concurrent validity of the BDI-II and the BDI-IIR, Pearson product moment correlation coefficients were calculated between total scores of each measure and the CES-D and SRS respectively. As seen in Table 5, the BDI-II was found to correlate .72 with the CES-D at Time 1 and .82 at Time 2. A test for the difference between dependent correlations (Bruning & Kintz, 1987) was performed and the correlations were found to differ significantly from Time 1 to Time 2, $t(42) = 2.14$, $p < .05$. The BDI-IIR was found to correlate .55 with the CES-D at Time 1 and .67 at Time 2. The correlations did not differ significantly across time, $t(45) = 1.54$. A test for the difference between independent correlations (Fisher's z; Bruning& Kintz, 1987) was performed to compare the strength of correlation between the BDI-II and the BDI-IIR with the CES-D. The two versions of the BDI-II did not differ significantly in their correlation with the CES-D at either Time 1 ($z = 1.33$) or Time 2 ($z = 1.56$).

The procedure recommended by Bruning and Kintz (1987, p.228) was used to test the difference between dependent correlations. The BDI-II correlated .72 with the SRS at Time 1 and .85 at Time 2. The correlations differed significantly across time, $t(42) = 3.01$, $p < .05$. The random version correlated .70 with the SRS at Time 1 and .77 at Time 2. No significant differences were found across time, $t(46) = 1.10$. Furthermore, the BDI-II and its random

Table 5

Alpha Coefficients and Intercorrelations of Measures of Depression and Social Desirability for the Total Sample (N = 139)[a]

Scale	1	2	3	4	5	6	7	8	9	10
1. BDI-II(1)	87									
2. BDI-IIR(1)	--[b]	85								
3. CES-D(1)	72	55	90							
4. SRS(1)	72	70	70	82						
5. ESDS(1)	-59	-77	-58	-72	81					
6. BDI-II(2)	86	58	39	54	-70	89				
7. BDI-IIR(2)	62	75	78	67	-64	--	93			
8. CES-D(2)	73	64	57	62	-61	82	67	92		
9. SRS(2)	64	61	59	70	-70	85	77	76	86	
10.ESDS(2)	-68	-71	-57	-65	87	-79	-73	-62	-79	83

Note. [a] Alpha coefficients are presented on the diagonal. All coefficients are significant at at least $p < .05$. Decimals are omitted. The number in brackets indicates the testing session.

[b] Indicates that a coefficient could not be computed.

counterpart did not differ significantly in their relation with the SRS at either Time 1 (z = .08) or Time 2 (z = 1.10).

In order to further evaluate the concurrent validity of the BDI-II and BDI-IIR, an aggregate measure of depression was formed to permit a more representative evaluation of convergent validity. Total scores on the CES-D and SRS were transformed into z-scores and subsequently amalgamated to form a composite depression index. The composite was found to correlate .78 with the BDI-II and .66 with the BDI-IIR at Time 1. The composite did not differ significantly in regards to its relationship with the BDI-II and its random counterpart (z = 1.18), indicating that the two versions of the BDI-II possess similar concurrent validity with other measures of depression. For Time 2, the BDI-II correlated .87 with the composite index whereas the BDI-IIR correlated .79. Again, no significant differences were found across measures. In addition, the BDI-IIR evidenced a significant difference in its correlation with the composite scale across time, t(45) = 2.04, p < .05, whereas the BDI-II did not, t(42) = .66.

To ascertain the test-retest reliability of the BDI-II and the BDI-IIR, Pearson product moment correlation coefficients were calculated separately for each measure. Test-retest reliability for the BDI-II was evaluated by examining the 13 participants who completed the BDI-II on

both testing occasions. As seen in Table 5, the reliability coefficient was estimated at .86. Test-retest reliability for the BDI-IIR was assessed in the same manner utilizing the 19 participants who completed the random version on both testing occasions, resulting in a test-retest coefficient of .75. The two versions of the BDI-II did not differ significantly in their test-retest reliability over a two week period (z = .79). Cronbach's alpha was calculated for the BDI-II and was found to be .87. The internal consistency of the random version was found to be similar at .85. Corrected item-total correlations for the BDI-II and the BDI-IIR are presented in Table 6.

To examine the relationship between social desirability and depression, the ESDS was correlated with the BDI-II and BDI-IIR respectively. The BDI-II was found to correlate -.59 with the ESDS at Time 1 and -.79 at Time 2. The correlation between social desirability and depression varied significantly over time, $t(42)$ = 4.17, p < .001. The random version of the BDI-II was found to correlate -.77 with the ESDS at Time 1 and -.73 at Time 2. No significant difference was found across time, $t(46)$ = .64. Upon comparison of the BDI-II with its random counterpart, no significant differences were found in regards to their relationship with the ESDS at either Time 1 (z = 1.53) or Time 2 (z = .64). Hence, the effects of social desirability do not differ across the two versions of the BDI-II.

Table 6

Corrected Item-Total Correlations of the BDI-II and BDI-IIR

Item	Symptom	BDI-II	BDI-IIR
1	Sadness	.57	.35
2	Pessimism	.41	.44
3	Past Failure	.32	.25
4	Loss of Pleasure*	.31	.66
5	Guilty Feelings	.49	.52
6	Punishment Feelings	.14	.51
7	Self-Dislike	.49	.50
8	Self-Criticalness	.48	.44
9	Suicidal Thoughts or Wishes	.13	.42
10	Crying	.33	.20
11	Agitation	.53	.19
12	Loss of Interest	.58	.57
13	Indecisiveness	.61	.40
14	Worthlessness	.55	.57
15	Loss of Energy	.55	.22
16	Changes in Sleeping Pattern	.60	.36
17	Irritability	.55	.63
18	Changes in Appetite	.56	.37
19	Concentration Difficulty	.60	.66
20	Tiredness or Fatigue	.56	.45
21	Loss of Interest in Sex	.25	.53

* Item-Total correlation on the BDI-II was significantly different from the Item-Total correlation on the BDI-IIR at $p < .05$.

Effects of Random Order

To compare the BDI-II and the BDI-IIR, one-way ANOVAs were performed for Time 1 and Time 2. At Time 1, the groups were collapsed and redivided into two groups: participants who completed the BDI-II versus those who completed the BDI-IIR. Hence, Groups 1, 4 and part of Group 7 represented the 46 participants in the BDI-II group at Time 1. The 47 participants in Groups 2, 3 and part of Group 7 comprised the Random group. Groups were redivided in the same manner at Time 2. No significant differences were found at Time 1, $F(1, 88) = .08$, $p = .78$, or at Time 2, $F(1, 89) = 1.39$, $p = .25$, indicating that the BDI-IIR did not result in significantly different total depression scores when compared to the BDI-II.

Pre-exposure Effects

To ascertain whether previous respondence results in a decrease in depression scores at later administrations, ANOVAs were performed on depression scores of participants who completed the same scale twice versus those who had not been previously exposed to a depression measure. The two groups in this case consisted of participants who completed the BDI-II on both occasions (Group 1) versus those in the control group who had not been previously exposed to a depression instrument (Group 5). The dependent variable was the total score on the BDI-II at Time 2. The two groups did not evidence a significant difference in BDI-II scores, $F(1,$

27) = .24, p = .63, indicating that previous completion of the BDI-II does not result in lower depression scores at a later administration. Similar findings were obtained for the BDI-IIR, $F(1, 32)$ = 2.47, p = .13.

To further evaluate the effects of exposure, within-subjects ANOVAs were performed using participants who completed the same measure at both Time 1 and Time 2. No significant differences were found for either the BDI-II, $F(1, 12)$ = .80, p = .39, or the BDI-IIR, $F(1, 18)$ = .20, p = .66. Thus, exposure did not reduce scores at the later administration.

In addition, within-subjects ANOVAs were performed on the two counterbalanced groups (Group 3 and Group 4) to ascertain whether the effects of exposure varied according to the presentation order of the measures. The first ANOVA was performed on group 3 and examined total scores on the BDI-IIR at Time 1 and total scores on the BDI-II at Time 2. No differences were found, $F(1, 14)$ = .04, p = .85. The second ANOVA examined group 4 scores obtained on the BDI-II at Time 1 and scores from the BDI-IIR at Time 2. Again, no differences were found, $F(1, 14)$ = .09, p = .77, indicating that exposure did not result in a decrease in later scores and that the presentation order of the measures did not affect subsequent depression scores.

Lastly, a two (high depression vs. low depression) by two (Time 1 vs. Time 2) mixed ANOVA was conducted on BDI-

II(R) scores to further evaluate the effects of exposure. A median-split was performed using the Time 1 composite index, whereby participants who scored above the median (median z = -0.01) were classified as high depression and those below the median were deemed low depression. No significant interaction was found, $F(1, 59)$ = .07, p = .79, indicating that high and low scorers did not experience differential patterns in scores between Time 1 and Time 2. Thus, exposure did not differentially affect later depression scores for depressed and non-depressed participants.

Discussion

The results of the present study provide information regarding the psychometric properties of the BDI-II and its random counterpart. In regards to concurrent validity, both versions of the BDI-II were found to exhibit strong correlations with the CES-D at Time 1 and Time 2. The BDI-II did not evidence a significantly stronger correlation with the CES-D when compared to its random counterpart, indicating that the two measures have similar concurrent validity. It was also noted that the concurrent validity for the BDI-II increased with the passage of time, whereas the BDI-IIR did not. Thus, although both the BDI-II and the BDI-IIR maintain an adequate level of convergent validity across time, it would appear that the former increases somewhat over time.

Similar findings emerged with regards to the SRS as

both BDI-II scales evidenced adequate concurrent validity. Again, the BDI-II evidenced a significant increase in convergent validity over a two-week interlude, whereas the random version did not. Nevertheless, no differences between measures were significant. Concurrent validity was also assessed through the aggregation of the CES-D and the SRS, allowing for a more stable and representative measure of depression. Both the BDI-II and its random counterpart evidenced adequate convergent validity with the aggregate measure at Time 1 and Time 2. Ergo, the two versions of the BDI-II maintain adequate concurrent validity after a two week interlude. Moreover, no significant differences were observed between measures. However, the BDI-IIR did evidence an increase in validity with the composite index across time. This difference was not observed for the BDI-II.

The observed relationship among the various measures with the passage of time is similar to the findings of Tanaka and Huba (1987) in which the BDI evidenced greater validity with both the CES-D and SRS after a one-month interval in college students. However, Tanaka and Huba (1987) did not offer an explanation for the observed increase in validity. Lambert, et al. (1986) and Lambert, et al. (1988), on the other hand, found that when the BDI and SRS were used to assess treatment effects, the measures became discrepant over time as they were speculated to be

assessing discordant symptomatology (Lambert et al., 1988).
This may be due to the use of a clinical sample and the
associated changes in depressive symptomatology and severity
resulting from treatment. As such, the results of the
present study, as well as those by Tanaka and Huba (1987),
may be more indicative of convergent validity in samples not
receiving treatment. Regardless, both the BDI-II and its
random counterpart possess good convergent validity.
Additionally, both versions of the BDI-II evidenced
convergent validity approximately equal to that of the
original BDI (Beck, et al., 1988), indicating that the new
version can be administered with confidence. Nevertheless,
the reason behind the increase in validity remains unclear.

The present study also examined the test-retest
reliability of the BDI-II and BDI-IIR. Reliability
coefficients over a two week period were moderate-to-high
for both measures and no differences were found between the
two versions of the BDI-II. Hence, both measures possess
similar test-retest reliability. The test-retest
reliability coefficients for the BDI-II and BDI-IIR are also
similar to those of the BDI. Thus, the BDI-II and BDI-IIR
evidence comparable test-retest reliability when compared to
their predecessor. Tanaka and Huba (1987), using the BDI,
estimated one-month test-retest coefficients to be .61,
considerably lower than the coefficients found by the
present study. When a two week interlude was examined,

Lightfoot and Oliver (1985) estimated the test-retest reliability of the BDI to be .90 in college students. Thus, it appears that the test-retest reliability of the BDI-II and its random counterpart in university populations are similar to that of their predecessor.

The results of the present study, then, indicate that the measures may be adequately employed in research where participants are screened for selection several days prior to the commencement of research or treatment implementation. Similarly, the results of the present study suggest that the measures can be employed in repeated monitoring of treatment efficacy. The strong test-retest reliability coefficients are advantageous as they indicate that individuals deemed as depressed at one testing are likely to score similarly at later testings, providing that the scores are the same. Thus, investigators who classify individuals prior to the commencement of a study need not be greatly concerned that measured depressive severity will change between the time of testing and the commencement of treatment, as has been suggested by Hatzenbuehler et al. (1983). Nevertheless, it would be beneficial to replicate these results with other non-psychiatric as well as clinical populations.

A third and related psychometric property examined by the present study was internal consistency. The mean internal consistency of the BDI in non-psychiatric populations has been estimated at .81 (Beck et al., 1988).

This is somewhat lower than the .87 and .85 estimates of internal consistency found for the BDI-II and BDI-IIR respectively in the present study. It would therefore appear that the addition of response options to the sleeping and eating items and/or the revision of the remaining items may have served to increase the internal consistency of the BDI-II. However, the findings of the present study are somewhat lower than the internal consistency of .93 found by Beck et al. (1996) using the new BDI-II with college students. Nevertheless, the difference is not significant and may not be a cause for concern.

The fourth psychometric property examined by the present study was discriminant validity. This form of construct validity was assessed through the relationship between the ESDS and the two versions of the BDI-II. Neither version evidenced adequate discriminant validity with the ESDS. Rather, the BDI-II and BDI-IIR exhibited strong negative correlations with social desirability. This correlation increased across time with regards to the BDI-II. Moreover, no differences were found in regards to the relationship of the BDI-II and BDI-IIR with the ESDS. In addition, the correlation between the two versions of the BDI-II and the ESDS are not discrepant from that of the BDI. When compared to the correlation between the BDI and the ESDS found in a student population by Cole (1988) and Tanaka-Matsumi and Kameoka (1986), -.65 and -.67

respectively, the results of the BDI-II and BDI-IIR found in the present study fare equivalently (see Table 5).

Of further interest is the comparison of total scores obtained using the BDI-II versus its random counterpart. Previous findings by Dahlstrom et al. (1990) evidenced a significantly higher mean depression score on the random version when compared to the original BDI. This finding was not replicated in the present study. The random version of the BDI-II did not evidence higher depression scores than the BDI-II. It may be the case that the higher scores found by Dahlstrom et al. (1990) on the BDI-R were anomalous in nature. Nevertheless, it is recommended that the present study be replicated with clinical samples. At present, the random version has only been tested on student populations.

As the two versions of the BDI-II have similar psychometric properties, it would appear that the measures are similar in nature and can therefore be used interchangeably. As such, there appears to be no advantage to utilizing the random version in place of the BDI-II.

The final purpose of the present study was to examine whether prior exposure to a depression instrument results in lower scores at a later administration. When the two control groups were compared to groups who had completed the same instrument on both testing occasions, namely groups 1 and 2, no significant differences were found. Participants who completed the BDI-II twice did not evidence diminished

depressive severity when compared to those who were not previously exposed to a depression scale. The same results were evidenced for the random version.

To further evaluate the effects of exposure, scores obtained by participants who completed the same measure twice were compared. If the speculative effects of exposure are correct, then scores obtained at Time 2 should be significantly lower than scores from Time 1. Again, no differences were found, indicating that exposure did not result in lowered scores at the later administration.

An additional comparison was performed whereby the effect of presentation order on later scores was examined. The two counterbalanced groups were examined to determine whether the specific version of the BDI-II presented at Time 1 would have differential consequences on the effects of exposure. That is, would the presentation of the BDI-II at Time 1 result in a greater exposure effect than the presentation of the BDI-IIR at Time 1, and vice-versa. Neither counterbalanced group evidenced a decrease in scores at the second testing session, again indicating that exposure does not result in decreased scores at later administrations. The lack of difference in scores also indicates that the effects of exposure do not vary based on which measure is presented first.

The final comparison pertaining to the effects of exposure was based on the hypothesis by Mark et al. (1991)

whereby participants deemed as depressed at Time 1 should evidence higher depression scores at Time 2 whereas non-depressed participants will evidence a decrease in scores at Time 2. No difference between high and low scorers was found across time, indicating that level of depressive severity at the time of initial testing does not affect scores obtained at later administrations. Again, exposure was not found to significantly influence depression scores obtained at later interludes.

As the present study examined students, it would be beneficial to compare the obtained results to those of other studies using students. Hatzenbuehler et al. (1983) and Hammen (1980) both examined college students and found significant decreases in BDI scores at later administrations. Similar results were found by Atkeson et al. (1982) in a clinical population. These findings are in opposition to those of the present study.

One explanation for this discrepancy is related to the composition of the present sample. Specifically, the sample of the present study may not be truly representative of college students in general. However, when compared to the characteristics of other college samples, the present sample would appear to be fairly similar in terms of depressive severity and social desirability. The mean across groups for the present sample on the BDI-II is similar to the BDI mean of 7.93 found by Dahlstrom et al. (1990). However, the

mean of the present sample is lower than the mean of 12.58 found with the BDI-II pilot college sample (Beck, et al., 1996). The mean score on the CES-D of the present sample is also similar to the mean of 15.51 found by Tanaka and Huba (1987), and scores on the ESDS are similar to those found in other college samples, 29.70 (Cole, 1988) and 30.21 (Holden & Fekken, 1989) respectively. It would therefore appear that the present sample is similar in regards to the CES-D and ESDS, but is slightly lower in depressive severity than the BDI-II pilot sample.

The present sample did differ with regards to the SRS. Previous studies examining college students have estimated the mean to be 33.20 (Cole, 1988) and 35.63 (Tanaka & Huba, 1987), which are considerably higher than the mean of 15.46 found with the current sample. Overall, however, the present sample appears representative of college students.

Although the results of the present study were discrepant from prior investigations involving the BDI and BDI-R, it does nevertheless proffer several strengths compared to previous studies. Specifically, the present study employed several measures of depression to ascertain the effectiveness of the random version of the BDI-II. The lack of difference in scores between the two versions of the Beck scale is further supported by a lack of difference in depressive severity found on the other measures of depression. Had a difference in severity been found without

the presence of additional measures of depression, it would remain uncertain whether the two groups were truly discrepant in regards to depressive severity or if the difference was spurious in nature. That is, the presence of the SRS, CES-D and composite scale permitted for a more complete assessment of the depressive severity of the BDI-II and BDI-IIR groups.

An additional strength pertains to the comprehensive manner in which the effect of exposure on later scores was evaluated. Previous studies employed either the same subjects at different intervals or compared scores of previous responders with control groups. The present study employed both methods and examined the effects of exposure within the same group of participants as well as across exposure versus no-exposure conditions. These methods permit the examination of exposure both within and across participants. The present study also examined whether scores obtained at later interludes vary according to depressive severity at the initial time of testing. Moreover, exposure was evaluated in the context of order of presentation, as was assessed with the counterbalanced groups. Hence, the effects of exposure were examined through a variety of methods, adding credence to the finding that exposure does not result in lower depression scores at later interludes.

The present study has also garnered knowledge regarding

the test-retest reliability, concurrent validity, discriminant validity, and internal consistency of the BDI-II and BDI-IIR. The present study has not only demonstrated the effectiveness of the BDI-II(R) with college students, but has found its psychometric properties to be similar to those of the original BDI. Additional strong points include the involvement of male participants, homogeneity of participants across groups, and the implementation of both control and counterbalanced groups.

In light of the newness of the BDI-II and its random counterpart, it would be beneficial for future studies to examine the psychometric properties of the BDI-II and the BDI-IIR in diverse clinical and non-clinical populations. This would not only increase knowledge regarding the efficacy of the measures, but would also improve the generalizability of the findings of the present study. Future studies would also benefit from the use of larger samples.

The results reported in the present study should, however, be interpreted with caution, given the limitations of the study. These include small sample sizes, skewed distributions, and a low range of depression scores, all of which may have affected the results. A concern associated with small sample size is that of low power. However, this may not be a great cause for concern as the means were in the opposite direction to those predicted for the pre-

exposure effect as well as the effects of randomization.   A further limitation pertains to the potentially high rate of Type I errors due to the testing of numerous hypothesis.

In summary, the present study has ascertained knowledge regarding several psychometric properties of the BDI-II and its random counterpart.   These properties are similar to those found in the BDI, and have in some cases been enhanced.   In addition, the present study failed to find a significant difference in depressive severity between the BDI-II and the random version, indicating that the two versions are similar in nature.   Lastly, prior exposure to a depression scale did not result in lower scores at a later interlude.

References

Abramson, L. Y., Seligman, M. E. P., & Teasdale, J. (1978). Learned helplessness in humans: Critique and reformulation. Journal of Abnormal Psychology, 87, 49-74.

American Psychiatric Association. (1994). Diagnostic and statistical manual of mental disorders (4th ed.). Washington, DC: Author.

Anastasi, A. (1988). Psychological testing (6th ed.). New York, NY: Macmillan.

Atkeson, B. M., Calhoun, K. S., Resick, P. A., & Ellis, E. M. (1982). Victims of rape: Repeated assessment of depressive symptoms. Journal of Consulting and Clinical Psychology, 50(1), 96-102.

Bech, P. (1992). Symptoms and assessment of depression. In E. S. Paykel (Ed.), Handbook of affective disorders (2nd ed.) (pp.3-14). New York, NY: The Guilford Press.

Beck, A. T. (1972). Measuring depression: The depression inventory. In T. A. Williams, M. M. Katz, & J. A. Shields (Eds.), Recent advances in the psychobiology of the depressive illnesses (pp. 299-302). Washington, DC: U.S. Government Printing Office.

Beck, A. T. (1976). Cognitive therapy and emotional disorders. New York: International University Press.

Beck, A. T., & Beamesderfer, A. (1974). Assessment of depression: The depression inventory. In P. Pichot (Ed.), Modern problems in pharmacopsychiatry (pp. 151-169). Basel,

Switzerland: Karger.

Beck, A. T., & Steer, R. A. (1987). Beck Depression Inventory manual. U.S.A.: The Psychological Corporation Harcourt Brace Jovanovich Inc.

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). Beck Depression Inventory II manual. San Antonio, TX: The Psychological Corporation Harcourt Brace Jovanovich Inc.

Beck, A. T., Steer, R. A., & Garbin, M. G. (1988). Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. Clinical Psychology Review, 8, 77-100.

Boyle, G. J. (1985). Self-report measures of depression: Some psychometric considerations. British Journal of Clinical Psychology, 24, 45-59.

Bruning, J. L., & Kintz, B. L. (1987). Computational handbook of statistics (3rd ed.) (pp. 226-229). Glenview, IL: Scott, Foresman & Co.

Butcher, J. N., Dahlstrom, W. G., Graham, J. F., Tellegen, A., & Kaemmer, B. (1989). Minnesota Multiphasic Personality Inventory-2 (MMPI-2): Manual for administration and scoring. Minneapolis: University of Minnesota Press.

Cambell, D. T., & Fisk, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. Psychological Bulletin, 50, 81-105.

Cole, D. A. (1988). Hopelessness, social desirability, and parasuicide in two college student samples. Journal of

Consulting and Clinical Psychology, 56(1), 131-136.

Crowne, D., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. Journal of Consulting Psychology, 24, 349-354.

Dahlstrom, W. G., Brooks, J. D., & Peterson, C. D. (1990). The Beck Depression Inventory: Item order and the impact of response sets. Journal of Personality Assessment, 55(1&2), 224-233.

Dick, W., & Hagerty, N. (1971). Topics in measurement. New York, NY: McGraw-Hill Book Company.

Edwards, A. L. (1957). The social desirability variable in personality assessment and research. New York, NY: The Dryden Press Inc.

Edwards, A. (1970). The measurement of personality traits by scales and inventories. New York: Holt, Rinehart & Winston.

Epstein, S. (1979). The stability of behavior: Predicting most of the people much of the time. Journal of Personality and Social Psychology, 37(7), 1097-1126.

Gabrys, J. B., & Peters, K. (1985). Reliability, discriminant and predictive validity of the Zung Self-Rating Depression Scale. Psychological Reports, 57, 1091-1096.

Gallagher, D., Nies, G., & Thompson, L. W. (1982). Reliability of the Beck Depression Inventory with older adults. Journal of Consulting and Clinical Psychology, 50(1), 152-153.

Gotlib, I. H., Lewinsohn, P. M., & Seeley, J. R. (1995). Symptoms versus a diagnosis of depression: Differences in psychosocial functioning. Journal of Consulting and Clinical Psychology, 63(1), 90-100.

Graziano, A. M., & Raulin, M. L. (1993). Research methods: A process of inquiry (2nd ed.). New York: Harper Collins.

Gregory, R. J. (1996). Psychological testing: History, principles, and applications (2nd ed.). Boston, MA: Allyn and Bacon.

Hamilton, M. (1960). A rating scale for depression. Journal of Neurology, Neurosurgery and Psychiatry, 23, 56-62.

Hammen, C. L. (1980). Depression in college students: Beyond the Beck Depression Inventory. Journal of Consulting and Clinical Psychology, 48(1), 126-128.

Hammen, C., Mayol, A., deMayo, R., & Marks, T. (1986). Initial symptom levels and the life-event-depression relationship. Journal of Abnormal Psychology, 95(2), 114-122.

Hathaway, S. R., & McKinley, J. C. (1942). A multiphasic personality schedule (Minnesota): III. The measurement of symptomatic depression. Journal of Psychology, 4, 73-84.

Hatzenbuehler, L. C., Parpal, M., & Matthews, L.

(1983). Classifying college students as depressed or nondepressed using the Beck Depression Inventory: An empirical analysis. Journal of Consulting and Clinical Psychology, 51(3), 360-366.

Holden, R. R., & Fekken, C. (1989). Three common social desirability scales: Friends, acquaintances, or strangers? Journal of Research in Personality, 23, 180-191.

Holden, R., Mendonca, J., & Mazmanian, D. (1985). Relation of response set to observed suicide intent. Canadian Journal of Behaviourial Science, 17, 359-367.

Jackson, D. N. (1989). Personality Research Form Manual (3rd ed.). Port Huron, MI: Sigma Assessment Systems.

Joseph, S., Lewis, C. A., & Olsen, C. (1996). Convergent validity of the Depression-Happiness Scale with measures of depression. Journal of Clinical Psychology, 52(5), 551-554.

Katz, R., Shaw, B., Vallis, T. M., & Kaiser, A. (1995). The assessment of severity and symptom patterns in depression. In E. Beckham & W. Leber (Eds.). Handbook of depression (2nd ed.). New York, NY: Godford Press.

Kerner, S. A., & Jacobs, K. W. (1983). Correlation between scores on the Beck Depression Inventory and the Zung Self-Rating Depression Scale. Psychological Reports, 53, 969-970.

Lambert, M. J., Hatch, D. R., Kingston, M. D., & Edwards, B. C. (1986). Zung, Beck, and Hamilton Rating

Scales as measures of treatment outcome: A meta-analytic comparison. Journal of Consulting and Clinical Psychology, 54(1), 54-59.

Lightfoot, S. L., & Oliver, J. M. (1985). The Beck Inventory: Psychometric properties in university students. Journal of Personality Assessment, 49(4), 434-436.

Linehan, M. M., & Nielsen, S. L. (1981). Assessment of suicide ideation and parasuicide: Hopelessness and social desirability. Journal of Consulting and Clinical Psychology, 49(5), 773-775.

Linehan, M. M., & Nielsen, S. L. (1983). Social desirability: Its relevance to the measurement of hopelessness and suicidal behaviour. Journal of Consulting and Clinical Psychology, 51(1), 141-143.

Margo, G. M., Dewan, M. J., Fisher, S., & Greenberg, R. P. (1992). Comparison of three depression rating scales. Perceptual and Motor Skills, 75, 144-146.

Mark, M. M., Sinclair, R. C., & Wellens, T. R. (1991). The effect of completing the Beck Depression Inventory on self-reported mood state: Contrast and Assimilation. Personality and Social Psychology Bulletin, 17(4), 457-465.

Mendonca, J. D., Holden, R. R., Mazmanian, D., & Dolan, J. (1983). The influence of response styles on the Beck Hopelessness Scale. Canadian Journal of Behavioral Sciences/Revue Canadienne du Science Compliquee, 15(3), 237-247.

Michelson, L., & Mavissakalian, M. (1983). Temporal stability of self-report measures in agoraphobia research. Behavioral Research and Therapy, 21(6), 695-698.

Miller, W. R., & Seligman, M. E. P. (1973). Depression and the perception of reinforcement. Journal of Abnormal Psychology, 82(1), 62-73.

Murphy, K. (1993). Psychology and the measurement game. Renfrew, ON: Spring Flower Press.

Murray, D. C. (1973). Suicidal and depressive feelings among college students. Psychological Reports, 33, 175-181.

Oliver, J. M., & Burkham, R. (1979). Depression in university students: Duration, relation to calendar time, prevalence, and demographic correlates. Journal of Abnormal Psychology, 88(6), 667-670.

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(6), 660-679.

Ossenkopp, K. P, & Mazmanian, D. S. (1985). The principle of aggregation in psychobiological correlational research: An example from the open-field test. Animal Learning and Behavior, 13(4), 339-344.

Pehm, L. P. (1976). Assessment of depression. In M. Hersen & A. S. Bellack (Eds.), Behavioral assessment: A practical handbook. Elmsford, NY: Pergamon Press.

Petrie, K., & Chamberlain, K. (1983). Hopelessness and

social desirability as moderator variables in predicting suicidal behaviour. Journal of Consulting and Clinical Psychology, 51(4), 485-487.

Radloff, L. S. (1977). The CES-D Scale: A self-report depression scale for research in the general population. Applied Psychological Measurement, 1(3), 385-401.

Santrock, J. W. (1993). Adolescence: An introduction (5th ed.). Dubuque, IA: Wm. C. Brown Communications, Inc.

Schnurr, R., Hoaken, P. C., & Jarrett, F. J. (1976). Comparison of depression inventories in a clinical population. Canadian Psychiatric Association Journal, 21, 473-476.

Strober, M., Green, J., & Carlson, G. (1981). Utility of the Beck Depression Inventory with psychiatrically hospitalized adolescents. Journal of Consulting and Clinical Psychology, 49(3), 482-483.

Tanaka, J. S., & Huba, G. J. (1987). Assessing the stability of depression in college students. Multivariate Behavioral Research, 22, 5-19.

Tanaka-Matsumi, J., & Kameoka, V. A. (1986). Reliabilities and concurrent validities of popular self-report measures of depression, anxiety, and social desirability. Journal of Consulting and Clinical Psychology, 54(3), 328-333.

Zung, W. W. K. (1965). A self-rating depression scale. Archives of General Psychiatry, 12, 63-70.

Appendix A

BDI-II
Student Number: _____   Age: ____
Date: _____   Sex: ____

This questionnaire consists of 21 groups of statements. Please read
each group of statements carefully, and then pick out the **one statement**
in each group that best describes the way you have been feeling during
the **past two weeks, including today.** Circle the number beside the
statement you have picked. If several statements in the group seem to
apply equally well, circle the highest number for that group. Be sure
that you do not choose more than one statement for any group, including
Item 16 (Changes in Sleeping Pattern) or Item 18 (Changes in Appetite)

1.Sadness
0   I do not feel sad.
1   I feel sad much of the time.
2   I am sad all the time.
3   I am so sad or unhappy that I
    can't stand it.

2.Pessimism
0   I am not discouraged about my
    future.
1   I feel more discouraged about my
    future than I used to be.
2   I do not expect things to work
    out for me.
3   I feel my future is hopeless and
    will only get worse.

3.Past Failure
0   I do not feel like a failure.
1   I have failed more than I should
    have.
2   As I look back, I see a lot of
    failures.
3   I feel I am a total failure as a
    person.

4.Loss of Pleasure
0   I get as much pleasure as I ever
    did from the things I enjoy.
1   I don't enjoy things as much as
    I used to.
2   I get very little pleasure from
    the things I used to enjoy.
3   I can't get any pleasure from
    the things I used to enjoy.

5.Guilty Feelings
0   I   don't   feel   particularly
    guilty.
1   I feel guilty over many things I
    have done or should have done.
2   I feel quite guilty most of the
    time.
3   I feel guilty all of the time.

6.Punishment Feelings
0   I don't feel I am being
    punished.
1   I feel I may be punished.
2   I expect to be punished.
3   I feel I am being punished.

7.Self-Dislike
0   I feel the same about myself as
    ever.
1   I have lost confidence in
    myself.
2   I am disappointed in myself.
3   I dislike myself.

8.Self-Criticalness
0   I don't criticize or blame
    myself more than usual.
1   I am more critical of myself
    than I used to be.
2   I criticize myself for all of my
    faults.
3   I blame myself for everything
    bad that happens.

9.Suicidal Thoughts or Wishes
0   I don't have any thoughts of
    killing myself.
1   I have thoughts of killing
    myself, but I would not carry
    them out.
2   I would like to kill myself.
3   I would kill myself if I had the
    chance.

10. Crying
0   I don't cry anymore than I used
    to.
1   I cry more than I used to.
2   I cry over every little thing.
3   I feel like crying, but I can't.

____   Subtotal Page 1

11.Agitation

0   I am no more restless or wound up than usual.
1   I feel more restless or wound up than usual.
2   I am so restless or agitated that it's hard to stay still.
3   I am so restless or agitated that I have to keep moving or doing something.

12.Loss of Interest

0   I have not lost interest in other people or activities.
1   I am less interested in other people or things than before.
2   I have lost most of my interest in other people or things.
3   It's hard to get interested in anything.

13.Indecisiveness

0   I make decisions about as well as ever.
1   I find it more difficult to make decisions than usual.
2   I have much greater difficulty making decisions than I used to.
3   I have trouble making any decisions.

14.Worthlessness

0   I do not feel I am worthless.
1   I don't consider myself as worthwhile and useful as I used to.
2   I feel more worthless as compared to other people.
3   I feel utterly worthless.

15.Loss of Energy

0   I have as much energy as ever.
1   I have less energy than I used to have.
2   I don't have enough energy to do very much.
3   I don't have enough energy to do anything.

16.Changes in Sleeping Pattern

0   I have not experienced any change in my sleeping pattern.
1a I sleep somewhat more than usual.
1b I sleep somewhat less than usual.
2a I sleep a lot more than usual.
2b I sleep a lot less than usual.
3a I sleep most of the day.
3b I wake up 1-2 hours early and can't get back to sleep.

17.Irritability

0   I am no more irritable than usual.
1   I am more irritable than usual.
2   I am much more irritable than usual.
3   I am irritable all the time.

18.Changes in Appetite

0   I have not experienced any change in my appetite.
1a My appetite is somewhat less than usual.
1b My appetite is much greater than usual.
2a My appetite is much less than before.
2b My appetite is much greater than usual.
3a I have no appetite at all.
3b I crave food all the time.

19.Concentration Difficulty

0   I can concentrate as well as ever.
1   I can't concentrate as well as usual.
2   It's hard to keep my mind on anything for very long.
3   I find I can't concentrate on anything.

20.Tiredness or Fatigue

0   I am no more tired or fatigued than usual.
1   I get more tired or fatigued more easily than usual.
2   I am too tired or fatigued to do a lot of the things I used to do.
3   I am too tired or fatigued to do most of the things I used to do.

21.Loss of Interest in Sex

0   I have not noticed any recent change in my interest in sex.
1   I am less interested in sex than I used to be.
2   I am much less interested in sex now.
3   I have lost interest in sex completely.

_____ Subtotal Page 2
_____ Subtotal Page 1

_____ Total Score

BDI-IIR

Student Number: _____   Age: ____
Date: _____   Sex: ____

This questionnaire consists of 21 groups of statements. Please read each group of statements carefully, and then pick out the **one statement** in each group that best describes the way you have been feeling during the **past two weeks, including today**. Circle the number beside the statement you have picked. If several statements in the group seem to apply equally well, circle the one that best describes you. Be sure that you do not choose more than one statement for any group, including Item 16 (Changes in Sleeping Pattern) or Item 18 (Changes in Appetite).

1. **Sadness**
   0 I am so sad or unhappy that I can't stand it.
   1 I do not feel sad.
   2 I am sad all the time.
   3 I feel sad much of the time.

2. **Pessimism**
   0 I am not discouraged about my future.
   1 I feel more discouraged about my future than I used to be.
   2 I do not expect things to work out for me.
   3 I feel my future is hopeless and will only get worse.

3. **Past Failure**
   0 I do not feel like a failure.
   1 I feel I am a total failure as a person.
   2 I have failed more than I should have.
   3 As I look back, I see a lot of failures.

4. **Loss of Pleasure**
   0 I get very little pleasure from the things I used to enjoy.
   1 I get as much pleasure as I ever did from the things I enjoy.
   2 I don't enjoy things as much as I used to.
   3 I can't get any pleasure from the things I used to enjoy.

5. **Guilty Feelings**
   0 I feel quite guilty most of the time.
   1 I don't feel particularly guilty.
   2 I feel guilty over many things I have done or should have done.
   3 I feel guilty all of the time.

6. **Punishment Feelings**
   0 I feel I may be punished.
   1 I don't feel I am being punished.
   2 I feel I am being punished.
   3 I expect to be punished.

7. **Self-Dislike**
   0 I feel the same about myself as ever.
   1 I dislike myself.
   2 I have lost confidence in myself.
   3 I am disappointed in myself.

8. **Self-Criticalness**
   0 I am more critical of myself than I used to be.
   1 I don't criticize or blame myself more than usual.
   2 I blame myself for everything bad that happens.
   3 I criticize myself for all of my faults.

9. **Suicidal Thoughts or Wishes**
   0 I would kill myself if I had the chance.
   1 I don't have any thoughts of killing myself.
   2 I have thoughts of killing myself, but I would not carry them out.
   3 I would like to kill myself.

10. **Crying**
   0 I cry more than I used to.
   1 I feel like crying, but I can't.
   2 I don't cry anymore than I used to.
   3 I cry over every little thing.

____ Subtotal Page 1

11. Agitation

0 I am so restless or agitated that I have to keep moving or doing something.
1 I am no more restless or wound up than usual.
2 I am so restless or agitated that it's hard to stay still.
3 I feel more restless or wound up than usual.

12. Loss of Interest

0 I have lost most of my interest in other people or things.
1 I have not lost interest in other people or activities.
2 I am less interested in other people or things than before.
3 It's hard to get interested in anything.

13. Indecisiveness

0 I find it more difficult to make decisions than usual.
1 I make decisions about as well as ever.
2 I have much greater difficulty making decisions than I used to.
3 I have trouble making any decisions.

14. Worthlessness

0 I feel more worthless as compared to other people.
1 I do not feel I am worthless.
2 I don't consider myself as worthwhile and useful as I used to.
3 I feel utterly worthless.

15. Loss of Energy

0 I have less energy than I used to have.
1 I don't have enough energy to do very much.
2 I don't have enough energy to do anything.
3 I have as much energy as ever.

16. Changes in Sleeping Pattern

0 I have not experienced any change in my sleeping pattern.
1a I sleep a lot more than usual.
1b I sleep a lot less than usual.
2a I sleep somewhat more than usual.
2b I sleep somewhat less than usual.
3a I sleep most of the day.
3b I wake up 1-2 hours early and can't get back to sleep.

17. Irritability

0 I am much more irritable than usual.
1 I am irritable all the time.
2 I am more irritable than usual.
3 I am no more irritable than usual.

18. Changes in Appetite

0 My appetite is much less than before.
   My appetite is much greater than before.
1a I have no appetite at all.
1b I crave food all the time.
2a My appetite is somewhat less than usual.
2b My appetite is somewhat greater than usual.
3 I have not experienced any change in my appetite.

19. Concentration Difficulty

0 I find I can't concentrate on anything.
1 I can't concentrate as well as usual.
2 I can concentrate as well as ever.
3 It's hard to keep my mind on anything for very long.

20. Tiredness or Fatigue

0 I am no more tired or fatigued than usual.
1 I am too tired or fatigued to do most of the things I used to do.
2 I am too tired or fatigued to do a lot of the things I used to do.
3 I get more tired or fatigued more easily than usual.

21. Loss of Interest in Sex

0 I have not noticed any recent change in my interest in sex.
1 I am less interested in sex than I used to be.
2 I have lost interest in sex completely.
3 I am much less interested in sex now.

____ Subtotal Page 2
____ Subtotal Page 1

____ Total Score

Appendix C

Order of Items for the Random Beck Depression Inventory II

Item	Order
1	3, 0, 2, 1
2	0, 1, 2, 3
3	0, 3, 1, 2
4	2, 0, 1, 3
5	2, 0, 1, 3
6	1, 0, 3, 2
7	0, 3, 1, 2
8	1, 0, 3, 2
9	3, 0, 1, 2
10	1, 3, 0, 2
11	3, 0, 2, 1
12	2, 0, 1, 3
13	1, 0, 2, 3
14	2, 0, 1, 3
15	1, 2, 3, 0
16	0, 2a, 2b, 1a, 1b, 3a, 3b
17	2, 3, 1, 0
18	2a, 2b, 3a, 3b, 1a, 1b, 0
19	3, 1, 0, 2
20	0, 3, 2, 1
21	0, 1, 3, 2

Note. From Dahlstrom, Brooks, and Peterson, 1990.

Appendix D

Zung Self-Rating Scale

This questionnaire consists of 20 statements. Please read each statement carefully and indicate which category best describes the way you have felt during the **past two weeks**.

	A little of the Time	Some of the Time	Good Part of the Time	Most of the Time
1. I feel down-hearted and blue.	[ ]	[ ]	[ ]	[ ]
2. Morning is when I feel the best.	[ ]	[ ]	[ ]	[ ]
3. I have crying spells or feel like it.	[ ]	[ ]	[ ]	[ ]
4. I have trouble sleeping at night.	[ ]	[ ]	[ ]	[ ]
5. I eat as much as I used to.	[ ]	[ ]	[ ]	[ ]
6. I still enjoy sex.	[ ]	[ ]	[ ]	[ ]
7. I notice that I am losing weight.	[ ]	[ ]	[ ]	[ ]
8. I have trouble with constipation.	[ ]	[ ]	[ ]	[ ]
9. My heart beats faster than usual.	[ ]	[ ]	[ ]	[ ]
10. I get tired for no reason.	[ ]	[ ]	[ ]	[ ]
11. My mind is as clear as it used to be.	[ ]	[ ]	[ ]	[ ]
12. I find it easy to do the things I used to do.	[ ]	[ ]	[ ]	[ ]
13. I am restless and can't keep still.	[ ]	[ ]	[ ]	[ ]
14. I feel hopeful about the future.	[ ]	[ ]	[ ]	[ ]
15. I am more irritable than usual.	[ ]	[ ]	[ ]	[ ]
16. I find it easy to make decisions	[ ]	[ ]	[ ]	[ ]
17. I feel that I am useful and needed.	[ ]	[ ]	[ ]	[ ]
18. My life is pretty full.	[ ]	[ ]	[ ]	[ ]
19. I feel that others would be better off if I were dead.	[ ]	[ ]	[ ]	[ ]
20. I still enjoy the things I used to do.	[ ]	[ ]	[ ]	[ ]

Appendix E

Center For Epidemiological Studies-Depression

This questionnaire consists of 20 statements. Please read each
statement carefully and indicate which category best describes
the way you have been feeling during the past week.

A=Rarely or None of the Time (Less than 1 Day)
B=Some or a Little of the Time (1-2 Days)
C=Occasionally or a Moderate Amount of Time (3-4 Days)
D=Most or All of the Time (5-7 Days)

		A	B	C	D
1.	I was bothered by things that usually don't bother me.	[ ]	[ ]	[ ]	[ ]
2.	I did not feel like eating; my appetite was poor.	[ ]	[ ]	[ ]	[ ]
3.	I felt that I could not shake off the blues even with help from my family or friends.	[ ]	[ ]	[ ]	[ ]
4.	I felt that I was just as good as other people.	[ ]	[ ]	[ ]	[ ]
5.	I had trouble keeping my mind on what I was doing.	[ ]	[ ]	[ ]	[ ]
6.	I felt depressed.	[ ]	[ ]	[ ]	[ ]
7.	I felt that everything I did was an effort.	[ ]	[ ]	[ ]	[ ]
8.	I felt hopeful about the future.	[ ]	[ ]	[ ]	[ ]
9.	I thought my life had been a failure.	[ ]	[ ]	[ ]	[ ]
10.	I felt fearful.	[ ]	[ ]	[ ]	[ ]
11.	My sleep was restless.	[ ]	[ ]	[ ]	[ ]
12.	I was happy.	[ ]	[ ]	[ ]	[ ]
13.	I talked less than usual.	[ ]	[ ]	[ ]	[ ]
14.	I felt lonely.	[ ]	[ ]	[ ]	[ ]
15.	People were unfriendly.	[ ]	[ ]	[ ]	[ ]
16.	I enjoyed life.	[ ]	[ ]	[ ]	[ ]

17. I had crying spells.   [ ]   [ ]   [ ]   [ ]

18. I felt sad.   [ ]   [ ]   [ ]   [ ]

19. I felt that people dislike me.   [ ]   [ ]   [ ]   [ ]

20. I could not get "going".   [ ]   [ ]   [ ]   [ ]

Appendix F

Edwards Social Desirability Scale

This questionnaire consists of 39 statements. After reading each statement carefully, circle the response which best describes you. If the statement is true all of the time, or most of the time, circle **True**. If it is not true all of the time, or most of the time, circle **False**. Be sure to read each statement carefully and circle only one response per statement.

1. My hands and feet are usually warm enough.   True   False

2. I am very seldom troubled by constipation.   True   False

3. I find it hard to keep my mind on a task or job.   True   False

4. Most any time I would rather sit and daydream
    than to do anything else.   True   False

5. My family does not like the work I have chosen
    (or the work I intend to choose for my life work).   True   False

6. My sleep is fitful and disturbed.   True   False

7. I am liked by most people who know me.   True   False

8. I am happy most of the time.   True   False

9. Criticism or scolding hurts me terribly.   True   False

10. It makes me impatient to have people ask my
    advice or otherwise interrupt me when I am
    working on something important.   True   False

11. I have had periods in which I carried on
    activities without knowing later what I had
    been doing.   True   False

12. I cry easily.   True   False

13. I do not tire quickly.   True   False

14. I am not afraid to handle money.   True   False

15. It makes me uncomfortable to put on a stunt at
    a party even when others are doing the same sort
    of things.   True   False

16. I frequently notice my hand shakes when I try
    to do something.   True   False

17. It does not bother me particularly to see animals suffer.   True   False

18. I dream frequently about things that are best kept to myself.   True   False

19. My parents and family find more fault with me than they should.   True   False

20. I have reason for feeling jealous of one or more members of my family.   True   False

21. No one cares much what happens to you.   True   False

22. I usually expect to succeed in things I do.   True   False

23. I sweat very easily even on cool days.   True   False

24. When in a group of people I have trouble thinking of the right things to talk about.   True   False

25. I can easily make other people afraid of me, and sometimes do for the fun of it.   True   False

26. I am never happier than when alone.   True   False

27. Life is a strain for me much of the time.   True   False

28. I am easily embarrassed.   True   False

29. I cannot keep my mind on one thing.   True   False

30. I feel anxiety about something or someone almost all the time.   True   False

31. I have been afraid of things or people that I knew could not hurt me.   True   False

32. I am not unusually self-conscious.   True   False

33. People often disappoint me.   True   False

34. I feel hungry almost all the time.   True   False

35. I worry quite a bit over possible misfortunes.   True   False

36. It makes me nervous to have to wait.   True   False

37. I blush no more often than others.   True   False

38. I shrink from facing a crisis or difficulty.   True   False

39. I sometimes feel that I am about to go to pieces.   True   False

Appendix G

Demographics Questionnaire

Student Number: _____

Professor: _____

Class Section: _____

Age: _____

Sex: male/female

Are you currently trying to lose weight?   yes/no

Are you presently being treated for depression?   yes/no

Have you ever been diagnosed with a psychiatric disorder?

                                                                    yes/no

If yes, please list the disorder(s).

_____

_____

_____

THANK YOU FOR YOUR PARTICIPATION

Appendix H

Consumer Survey

Please answer the following questions as best as you can:

1-Which do you prefer: Coke   or   Pepsi

2-What is your favourite brand and flavour of soup:

_____

3-What is your favourite flavour of ice cream?

_____

4-Do you believe that "beer" should be advertised on
television, and why or why not?

_____

_____

_____

_____

5-Please describe what you like and dislike in both
television and radio advertisements.

_____

_____

_____

_____

_____

_____

6-When purchasing a product, what do you look for (ie-colour, ingredients, environmentally sound.....)?

_____

_____

_____

_____

_____

_____

7-Are there any products that you refuse to purchase. If yes, please explain why?

_____

_____

_____

_____

_____

THANK YOU FOR YOUR PARTICIPATION

Appendix I

Recruitment Speech

"Hi! My name is Beth Merkley and I'm working on my Masters degree in Clinical Psychology here at LU. I'm doing my thesis on mood and am here to recruit volunteers for my study. The study will involve completing a set of questionnaires during two separate sessions, with each session lasting approximately 25-40 minutes. Each session that you attend will be worth one bonus point, so if you come to both sessions, you'll receive two bonus points. All your responses will be confidential and you can withdraw from the study at any time if you so desire. In order to ensure confidentiality, I'll be using your student numbers and not your names. The student numbers will also ensure that you'll receive your bonus points and the numbers will be deleted when the data is entered into the computer. I'll be passing around a sign-up sheet during today's class, so if you're interested in participating, please sign up for a session. Don't forget to write down the time, date and location of the session you signed up for. Thanks lot and I hope to see you there."

Appendix J

Standardized Instructions-Time 1

"Hello! Thank you for volunteering. Your participation is greatly appreciated. As I explained in class, you will be given a package of questionnaires to answer. You may withdraw from the study at any time should you so choose, but in order to receive your second bonus credit, you will have to return to answer another package of questionnaires in two weeks. You can sign up for the next session once you have finished today's questionnaires. It will take approximately 25-40 minutes per session to answer the questionnaires. Please remember to include your student number, class section and professor's name. These are for identification purposes only and will be destroyed once they are entered into the computer. All your responses will be kept confidential and I'll be visiting your classes in a few weeks to hand out a description of the study. Thank you again for your participation and if you have any questions, please do not hesitate to ask."

Standardized Instructions-Time 2

"Hello, welcome back. We will be doing the same thing as last time. I will hand you a set of questionnaires which will take approximately 25-40 minutes to complete. Again, if you wish to withdraw from the study, you may do so at any time. Please remember to include your student number, class section and professor's name. Just to remind you, this

information is for identification purposes only and will be destroyed when the data is entered into the computer. As well, all your answers will be kept confidential. Thanks again for participating, and if you have any questions, please don't hesitate to ask."

Appendix K

CONSENT FORM

My signature on this sheet indicates that I agree to
participate in a study by Beth Merkley (M.A student) under the
supervision of Dr. Dwight Mazmanian, on MOOD and also indicates
that I understand the following:

1. That I will be asked to complete a series of
    psychological questionnaires during two separate
    sessions separated by two weeks, and that each session
    will last approximately 25-40 minutes.
2. There are no known physical or psychological risks
    associated with participating in this experiment.
3. I am a volunteer and can withdraw from the study at
    any time without explanation and without penalty.
4. The data I provide will be confidential.
5. I will receive a summary of the project following
    the completion of the project, if I so request.
6. The research data will be stored in a secured
    environment for seven years.

I have received explanations about the nature of the study,
its purpose, and procedures.

_____

Signature of Participant   Date

If you have any questions or concerns about this study,
please contact:
        Beth Merkley   343-8476   BB0029M
        Dr. Dwight Mazmanian   343-8257   SN1016

If participating in this study or completing the questionnaires has distressed you or has raised personal issues that you would like to discuss, or if you need someone to talk to, the following organizations are available:

RESOURCE NUMBERS

LU Health Centre   343-8361
Peer Support Line   343-TALK (8255)
Chaplain   343-8018
Psychology Clinic   343-8441
Native Support Services   343-8084
Career & Counselling Services   343-8018

Appendix L

RESEARCH SUMMARY

Thank you for your participation in this study on MOOD. The purpose of the study is to examine the effects of re-ordering the presentation of a series of questions pertaining to mood. During the study, you completed either the original version of a psychological scale measuring mood and/or the re-ordered version of the same scale in which the response items were presented in an irregular manner. A past study by Dahlstrom, Brooks and Peterson (1990) found that individuals who completed the re-ordered version of the scale scored differently than those who completed the original version. The present study has sought to replicate these results, and to determine if the way you respond to the questionnaire(s) changes over time. That is, did your responses change the second time you completed the questionnaire.

In order to test these two hypotheses, approximately half of the participants completed the original scale and the other half completed the re-ordered version during each of the two testing sessions (the first session and again two weeks later). Some participants completed the same version twice, while others completed a different version each time. This will help determine if the responses vary according to the specific form of the questionnaire. In addition, some participants completed a consumer survey rather than a mood questionnaire during the first session. The consumer survey was a distracter task, and will determine if people who complete a mood questionnaire twice respond differently at the time of the second session when compared to those who have not previously completed a mood questionnaire. Just an added point of interest, Coke was overwhelmingly preferred to Pepsi and your favourite brand of soup was Campbell's Chicken Noodle, followed closely by Lipton Chicken Noodle.

If you are interested in obtaining a summary of the results, please write your permanent address on one of the available address labels. We will mail the summary to you after the data analyses are complete. This may be several months after you participated in the study. Thanks again for participating.
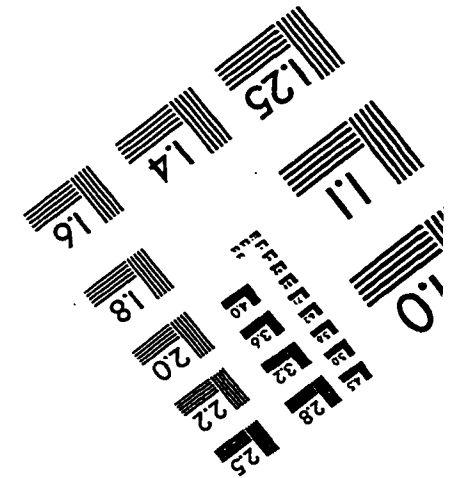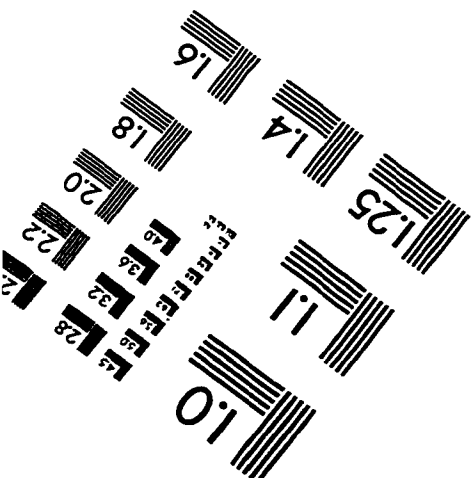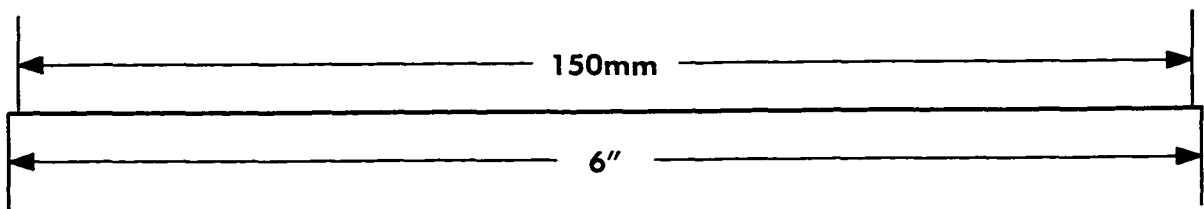
Sincerely,

Beth Merkley
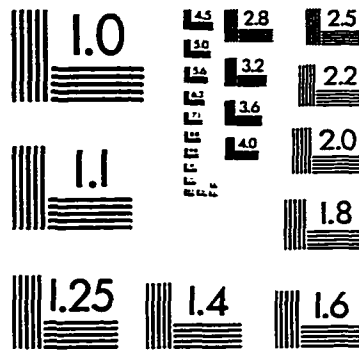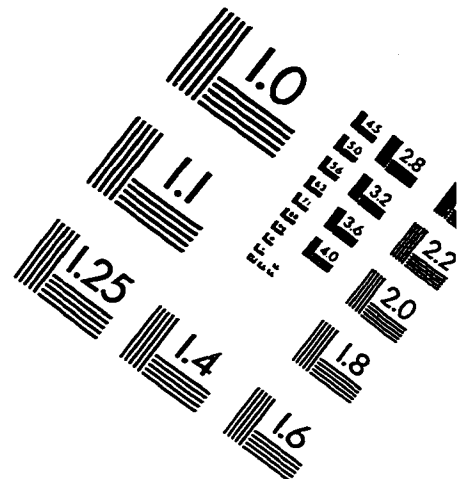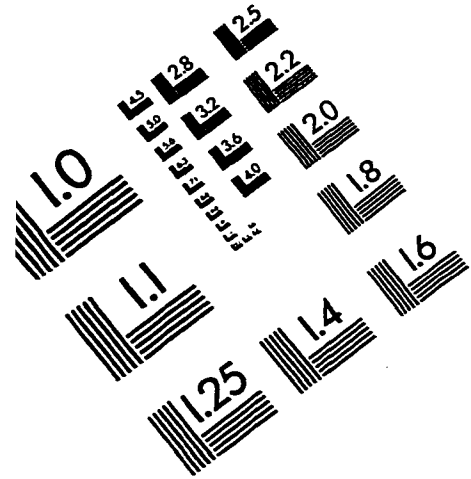
Contact Numbers:
    Beth Merkley   343-8476 BB0029M
    Dr. Mazmanian   343-8257 SN1016

Reference
    Dahlstrom, W. G., Brooks, J. D., & Peterson, C. D. (1990). The Beck Depression Inventory: Item order and the impact of response sets. Journal of Personality Assessment, 55(1&2), 224-233.

IMAGE EVALUATION
TEST TARGET (QA-3)

150mm

6"

APPLIED IMAGE . Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989