A Comparison Of The Vulnerability Of The Beck Depression Inventory And The Modified Stroop Procedure To Intentional Response Alteration

Corresponding Author:
Dr. Eugene Lubliner,
Assistant Adjunct Professor, Hofstra University, Hauser Hall, Room 220, Hofstra University, 11549 - United States of America

Submitting Author:
Dr. Eugene P Lubliner,
Assistant Adjunct Professor, Hofstra University, Hauser Hall, Room 220, Hofstra University, 11549 - United States of America

Article ID: WMC001366
Article Type: Research articles
Submitted on: 07-Jan-2011, 02:48:58 PM GMT   Published on: 08-Jan-2011, 07:34:59 PM GMT
Article URL: http://www.webmedcentral.com/article_view/1366
Subject Categories: PSYCHOLOGY
Keywords: Beck Depression Inventory, Modified Stroop Procedure, Stroop, Depression, Response Alteration, Lying

How to cite the article: Lubliner E, Motta R. A Comparison Of The Vulnerability Of The Beck Depression Inventory And The Modified Stroop Procedure To Intentional Response Alteration. WebmedCentral

Source(s) of Funding:
None.

Competing Interests:
None.
A Comparison Of The Vulnerability Of The Beck Depression Inventory And The Modified Stroop Procedure To Intentional Response Alteration

Author(s): Lubliner E, Motta R

Abstract

The Beck Depression Inventory – Second Edition has been validated in numerous populations and is the most widely used depression assessment tool in research and in practice. The obvious nature of its items makes the instrument vulnerable to intentional efforts to distort one’s level of depression. The purpose of this study was to examine the relative resistance to intentional response distortion on the BDI-II in comparison to a modified Stroop procedure involving the color naming of depression relevant words. Mildly depressed and non-depressed participants were asked to present themselves as either depressed or non-depressed on each of the two instruments. Each participant was provided with a vignette that encouraged them to respond in a depressed or non-depressed manner. Thus, the study involved a 2 x 2 design with two levels of affect (depressed and non-depressed) and two levels of response distortion (in a depressed or non-depressed direction). Study results indicated that both depressed and non-depressed participants were able to intentionally alter their responses to appear more or less depressed on the BDI-II but were unable to do this on the modified Stroop.

Introduction

The Beck Depression Inventory – Second Edition1 is a popular self-report measure of depression. The current version of the inventory and its predecessors has been shown to be reliable and valid across a wide age range of samples.1 Additionally, the BDI-II and previous versions have been standardized in different countries and cultures. It has been shown to be reliable and valid in Bulgaria2, Spain3, Cuba4, China5, Japan6, and many subcultures.7-9 Despite the volume of research supporting the use of the BDI-II, there are certain instances when the inventory might be insufficient or inaccurate in assessing a person’s level of depression. One such scenario is when a researcher is interested in studying trait depression, as opposed to state depression. The BDI-II is sensitive to minor fluctuations in a person’s mood.10,11 Therefore, if one wishes to examine trait depression, it may be an ineffective measure. Another situation in which the BDI-II may be ineffective is when the person being assessed is motivated to alter the depression level of their responses.12,13 Throughout the years many questions have been raised regarding a person’s ability to fabricate responses on the BDI-II.12-15 A person might intentionally distort responses to appear to have emotional difficulties or deny these difficulties altogether.

In these specific instances that weaken the validity of the BDI-II, an alternative measure might be used. A measure such as the modified Stroop may be less susceptible to attempts to give preferred responses. The modified Stroop procedure, as used here, involves color-naming depression related words. The person who is depressed will take longer to name the color of a depression related word because the word itself is troubling and therefore interferes with color naming. This delay in color naming is known as the Stroop interference effect. In terms of Beck’s model of depression, the modified Stroop indirectly assesses a person’s cognitive schema for depression. If a person has a pervasive schema for depression, then depression related words would cause high amounts of disturbance on a Stroop task.

There might be a number of testing situations where a person might wish to appear more or less depressed. If this were the case, it would be relatively easy to alter one’s responses on the BDI-II as the items are relatively obvious with regard to what constitutes a depressed or non-depressed response. In contrast, it would be far more difficult to alter one’s responses on the Stroop as the task is simply one of response latency in color naming and it is not obvious as to how one should respond so as to appear more or less depressed.

While the Stroop appears to be far more resistant than the BDI-II to deliberate response alteration, it does not have the cutoff scores that the BDI-II has and therefore is relatively lacking in incremental validity. When using the Stroop to identify depressed responding, one cannot state that a given response time represents, say, mild or moderate depression. However, initial findings from the current study do
show significant correlations between the Stroop response times and known measures of depression (i.e. the Revised Hamilton Rating Scale for Depression16 and the Zung Self-Rating Depression Scale17, r = .39 - .49, p < .05). The purpose of this study is to determine how resistant the BDI-II and the modified Stroop task are toward attempts at manipulating one’s responses. This will be accomplished by creating situations where one is motivated to fabricate responses. Participants who have elevated levels of depression and non-depressed controls will be identified. These participants will then be read vignettes that encourage response distortion to appear depressed or non-depressed and comparisons between scores on BDI-II and latency times on the modified Stroop will be examined.

Overview of the BDI-II

Beck18 theorized that depressed individuals have biased schemata—meaning they encode, store, and retrieve information differently than a non-depressed individual. This bias leads to the misinterpretation of events and actions. Strengths of the BDI include its ease of use, its applicability to diverse samples, and the fact that it has been the subject of numerous studies since its creation. Research has shown that the BDI is able to consistently and accurately measure current levels of depression in many different settings. Meta-analyses that examine the BDI have shown that the instrument has an overall test-retest reliability of 0.72.19 Other research that examines the internal consistency of the BDI shows that all items on the inventory load onto one overall factor that can be characterized as general depression.20 These studies support the BDI as a valid instrument for assessing depression.

The BDI is easy to use.21 It is convenient to have a short and easily administered assessment tool that can reliably predict symptoms of depression and can be completed with only minimal involvement from a researcher or mental health professional. By simply reading through the test items of the BDI, it is clear that the inventory has a high degree of face validity. However, ease of administration might be outweighed by the potential to incorrectly classify someone as depressed or non-depressed.

The major critiques of the BDI-II can be categorized in the following groups: lack of specificity, sensitivity to minor mood fluctuations, and demand characteristics that confound results. One of the greatest strengths of the BDI-II is also its greatest weakness: it is sensitive to slight changes in response patterns. These changes could be caused by fluctuating levels of overall depression, but changes could also be caused by other psychopathologies aside from depression, by temporal changes in current mood that do not speak to more general mood patterns, or by willfully manipulating answers to achieve a desired result. The BDI has been shown to lack low-12 and high-end specificity22. Low-end specificity is the ability of an instrument to correctly identify people exhibiting low levels of a certain behavior by scoring low on the instrument. High-end specificity is the opposite, where an instrument can correctly identify people with high levels of a certain behavior by scoring high on the instrument. Factors such as malingering, denial, and overall level of psychopathology affect how a person answers the BDI-II at the high and low ends.

Test-retest reliability of the BDI is affected when a person is read a depression-related vignette in between testing sessions.23 The level of depression at the second testing, as assessed by the BDI, increased based on the level of depression of the vignette. This demonstrates that the BDI is sensitive to less severe depression.13 Other studies have examined organization of the BDI as a demand characteristic. The clearly labeled titles of each section, along with the ordered item responses that go from least depressed symptoms to most depressed symptoms, leave little to the imagination. One study demonstrated that when the response items of the BDI are randomly presented, more accurate results were obtained.14 This random presentation seems to be able to counteract some of the demand characteristics inherent in the current presentation of the questionnaire.

These weaknesses of the BDI-II might be addressed by using a different measure to assess depression. One such assessment measure is the modified Stroop procedure.

Overview of the Modified Stroop Procedure

The modified Stroop (or emotional Stroop) task is an alternative version of the classic Stroop procedure.24 This task uses emotionally charged words printed in different colored inks as the stimulus. Words that have an emotional significance to the person performing the
task are likely to cause more interference than they would to a person who is emotionally unaffected by the words. This will result in longer color naming times for the people experiencing greater interference. The modified Stroop task has been shown to be able to successfully discriminate between people with depression versus people without depression.25-30

Less consistent findings have also been reported.31-33

The major strength of the modified Stroop procedure is its relative stability over time. Research suggests that the modified Stroop for depression is consistent over many months between testing intervals.34 This stability is what provides the basis for the hypothesis that the modified Stroop may be resistant to attempts at fabricating responses. Since the items in a Stroop task are novel to most people, the Stroop is resistant to response manipulation attempts.

Research has demonstrated that the modified Stroop procedure more consistently discriminates between depressed and non-depressed people than the BDI over the span of a year.34 This implies that the modified Stroop assesses more stable trait depression, whereas the BDI may be affected by transient state depression. Longitudinal studies have shown that depressed psychiatric patients show a decrease in Stroop interference times after a psychotherapeutic intervention.29,35 If the modified Stroop measures trait depression, then it seems reasonable that Stroop times decrease after an intervention for depression. If a person’s overall level of depression decreases, this will in turn create less interference on the Stroop task. In a modified Stroop procedure, response delays should only be the result of cognitive disruption caused by inadvertently reading an emotionally charged word while trying to name its color. Some other reasons that word items can delay response latency are because they have longer syllable length, are used less frequently in the English language, or are less concrete and salient. Unfortunately, emotionally charged words have a higher probability of being longer in length, used less frequently, and being less concrete and salient.36 Therefore, careful selection of neutral words must be considered in any experiment. They have to meet the characteristics of the selected emotional words.

It also may be possible to create longer interference times within the neutral word items. By increasing the amount of emotional disruption a person experiences on the neutral words one would eliminate any observable Stroop effect, because there would be little or no time difference between neutral and emotionally charged items. When neutral words and emotionally charged words are presented in a mixed presentation, carryover effects from reading an emotionally charged word can spill over to delay the color naming decision on neutral words.37,38 This is why current research suggests the best way to evaluate a modified Stroop effect is to present items in block form.39,40 In one unusual example, researchers were able to create a reverse modified Stroop effect by ordering items in a pseudo-random pattern where an emotionally charged word is always followed by two neutral words. When this pattern was used, color naming for neutral words was significantly longer than color naming for emotionally charged words.37

Another critique of the modified Stroop procedure is that there is a paucity of knowledge of the underlying cognitive mechanisms that explain the modified Stroop procedure. The original belief held by most researchers investigating the modified Stroop was that the cognitive mechanisms behind this phenomenon were the same as those in the classic Stroop procedure. The classic Stroop procedure uses items that are semantically related to each other (i.e. color words and actual ink color). Alternately, the modified Stroop uses items that are not semantically related to color. There is no ink color that is logically associated with the word “failure.” Current research shows that a classic Stroop task involves immediate cognitive disruption due to processing two semantically related items simultaneously.39 The modified Stroop procedure appears to involve a slowing of response times caused by the word item incorporating itself into a person’s cognitive schema.37-39 Thus, when a word like “failure” is printed in green, it takes longer for a depressed person to say, “green,” than if the underlying word is neutral such as “window.”

An additional critique of the Stroop procedure, as mentioned earlier is its lack of incremental validity. While significant relationships exist between paper and pencil measures of depression and Stroop color naming times of depression related words, cutoff response times have not been established that would specifically correspond to varying levels of depression. Thus, it appears that both the Stroop procedure and the BDI-II have strengths and weaknesses. The strength of the BDI-II is its ability to delineate levels of depression. Its weakness is the obvious nature of its items. The items on the Stroop are not easily seen as depression related because the task involves color naming; however, the measure lacks cutoff scores. One final critique of both the Stroop and BDI-II, and perhaps a number of similar assessment devices, is their relative lack of discriminant validity. It cannot be shown that the Stroop or the BDI-II assesses only depression and not other types of pathology. Given the overlap among different types of pathology and the
comorbidity that exists in most forms of psychological distress, it cannot be argued that these measures validly discriminate depression from other forms of psychological distress.

The present study investigates a participant’s ability to intentionally alter responses on the BDI-II and the modified Stroop task. Because there are numerous situations in which a person being assessed may be motivated to fabricate responses, the sensitivity of these two measures toward active attempts to distort one’s level of depressive affect is important.

Hypotheses Concerning Main Effects

Depressed participants will have significantly longer Stroop interference times than non-depressed participants on the modified Stroop card with depression related words, regardless of whether they were instructed to respond in a depressed or non-depressed manner. This is expected because the modified Stroop appears resistant to efforts to intentionally alter one’s responses.

Depressed and non-depressed participants will not differ from each other on the modified Stroop card with neutral words regardless of whether they were instructed to respond in a depressed or non-depressed manner. This is expected because the neutral words will not activate depressive cognitions or affect that may create additional task interference.

Participants who are told to respond in a depressed manner will have significantly higher scores on the BDI-II than participants who are told to respond in a non-depressed manner, regardless of whether they were initially depressed or initially non-depressed. This is because the BDI-II may be sensitive to attempts to deliberately alter one’s responses.

Hypotheses Within the Initial Level of Depression Condition

Within the initially depressed group, participants who are instructed to respond in a depressed and non-depressed manner will not differ from each other in Stroop response times for either the neutral or depression related word cards. Again, the Stroop is expected to be resistant to deliberate attempts at manipulating one’s responses.

Similarly, within the initially non-depressed group, participants who are instructed to respond in a depressed and non-depressed manner will not differ from each other in Stroop response times for either the neutral or depression related word cards.

Hypotheses Within the Vignette Condition

Within the group of participants encouraged to respond in a depressed manner, participants who are initially depressed and initially non-depressed will not differ from each other on BDI-II scores. This is expected because the BDI-II may be vulnerable to efforts of participants to manipulate their scores so as to appear depressed.

Similarly, within the group of participants encouraged to respond in a non-depressed manner, participants who are initially depressed and initially non-depressed will not differ from each other on BDI-II scores.

Method

Participants and Settings

There were 4 groups in this study, each of which contained 18 participants, for a total of 72 participants (36 depressed, 36 non-depressed). The participants were drawn from a community mental health center and a physical therapy rehabilitation facility. In order to be included in the study, participants had to meet certain criteria. A potential participant had to be a native speaker of English, could not be color-blind, and could not be diagnosed with a psychotic disorder or bipolar disorder. To be included in the depressed group, a participant had to be initially identified by the facility as having a psychiatric disorder, to be receiving treatment at the time of testing, but most importantly, had to score in at least the mildly depressed range on the two standardized depression diagnostic measures described below. In order to be included in the non-depressed group, a participant had to have no history of mental illness or had to be out of treatment for any psychiatric condition for at least one year and had to score in the non-depressed range on the two depression diagnostic measures.

Of the 36 participants in the depressed group, 29 were recruited from the community mental health center and 7 were recruited from the physical therapy rehabilitation facility. The mean age of the depressed group was 46.36 (SD = 16.19). There were a total of 10 males and 26 females. All of the 36 participants in the non-depressed group were recruited from the physical therapy rehabilitation facility. The mean age of the non-depressed group was 42.72 (SD = 16.30). There were a total of 19 males and 17 females.

Depression Diagnostic Measures

Revised Hamilton Rating Scale for Depression.41 The RHRS is a self-report instrument used to measure the severity of depressive symptoms in adults. The inventory has 76 items and has a check built into the instrument to assess item response validity and any inconsistency in responding. Each item is grouped into a set of depression related symptoms, such as “feelings of guilt,” “loss of insight,” or “weight loss.” The items in each group increase in severity of symptoms, and also increase in point value towards total depression score. The participant is asked to endorse...
negative predictive power (NPP). PPP means that a person who receives a high score on the SDS actually displays symptoms of depression based on clinical observation. NPP means that a person who receives a low score on the SDS does not display symptoms of depression. When SDS results were compared to actual clinician reports, the SDS had a PPP of .75 and an NPP of .68.44 This indicates that the SDS has a high level of diagnostic discrimination.

Dependent Measure

Beck Depression Inventory – Second Edition.1 The BDI-II is a self-report instrument used to measure the severity of depressive symptoms in adults and adolescents. The inventory has 21 items and can be used with participants from ages 13 and older. The item wordings and instructions were designed to be consistent with the DSM-IV criteria for major depression. The BDI-II utilizes a four-point scale ranging from 0 to 3 points. As each statement increases in point value, this reflects an increase in the severity of the depressive symptom. Scores on the BDI-II can range from 0 to 63. The sum of a participant's item scores generates an overall depression score which translates into four descriptive categories: 0 to 13 is minimal depression, 14 to 19 is mild depression, 20 to 28 is moderate depression, and 29 to 63 is severe depression.

The BDI-II has been shown to have a high internal consistency, with a coefficient alpha for a sample of 120 college students of .93, and a coefficient alpha for a sample of 277 psychiatric outpatients of .92.1 Additionally, the inventory has a high test-retest reliability for a sample of 26 psychiatric outpatients, r = .93, p < .01. The inventory has also been shown to have high convergent validity. Scores on the BDI-II are highly correlated with scores on the previous version of the inventory, r = .93, p < .01. The convergent validity was also demonstrated when the BDI-II was correlated with the Hamilton Rating Scale for Depression, r = .71, p < .05.16

Modified Stroop Procedure.24 The modified Stroop procedure is a modified version of the original Stroop task. In this version of the task, participants are asked to name the ink color of words with varying emotional valences, as opposed to color word names in the original task. For this task, participants are asked to name the color in which the word is printed while ignoring the actual content of the word.

For the current study, words were derived from Klieger and Cordner27, who used ten negative words and ten neutral words that were paired based on frequency using Kucera and Francis,45 length, and beginning letter. These 20 words were piloted by a random sample of 25 adults to determine the 5 most negative words. This was to ensure that the negative words
were the most relevant to the sample of similar age peers. When the five most negative words were determined, they were paired with the five neutral words previously established by Klieger and Cordner.27
The five negative words identified by a pilot study were "depressed," "failure," "gloomy," "hopeless," and "unhappy." The five neutral words selected by the pilot study were "disguised," "follows," "gamble," "harmless," and "utility." The five negative and neutral words were placed on individual cards (8.5 x 11). Additionally, a card was used that consisted of control stimuli (i.e., 00000). Each card had the 5 word items presented 20 times each, for a total of 100 items per card. Each word appeared four times in a given ink color (red, yellow, green, black, and blue), and the cards were arranged in such a way that neither the same word nor the same color were repeated next to, above, or below one another. Each participant received the cards in the following order: the control card, the neutral word card, and then the negative word card. This was done to ensure that any practice effect would go against the specific research hypotheses. By taking the time used to complete the control card and subtracting it from the time it took to complete the neutral and negative word cards, a modified Stroop interference time was calculated. The modified Stroop interference times for neutral and negative words were treated as two separate dependent variables during data analysis. Modified Stroop response time was calculated by subtracting the time to name the colors of items (colored zeros) on the control card from the times to color name items on the neutral or negative cards.

Procedure
Participants at the two sites mentioned above signed informed consent forms and indicated whether or not they were in treatment for a psychological disorder. They also indicated whether or not they were color blind. In order to be categorized as depressed, participants had to fall in the mildly depressed range or above on both the SDS and RHRSD. In order to be categorized as non-depressed, participants had to fall in the non-depressed range on both the SDS and the RHRSD and could not be currently receiving psychiatric treatment. Participants who scored in the depressed range on one self-report measure and the non-depressed range on the other measure were not asked to take part in the second half of the study. After meeting all the initial criteria, the qualifying depressed and non-depressed participants were administered the second half of the study within one hour of completing the screening measures. Participants were randomly assigned into one of two conditions: either (1) being read a vignette that encouraged minimizing depressive symptoms (see Appendix A) or (2) being read a vignette that encouraged enhancing depressive symptoms (see Appendix B). Because the participants were all adults that were participating in some form of weekly therapeutic service, both vignettes focused on salient issues a client might encounter.

The participants were then instructed to answer the items of the two dependent measures as if they were experiencing the situation described in the vignette. The two dependent measures were presented to participants in a counterbalanced fashion, where half the participants in each of the four conditions received the BDI-II first and the other half received the modified Stroop first. Participants were given specific instructions for each of the measures immediately before they were presented. Before the participant received the BDI-II the instructions that appear on top of the BDI-II form were read aloud. Then the participant was given as much time as needed to complete the form.

When the participant was given the modified Stroop procedure, they were given the following instructions, “Starting on the top of the page, and going from left to right, please name the ink color of each item aloud without paying attention to the word itself. The word itself is of no concern, only focus on the ink color. Perform this task as quickly as you can without making any mistakes.” Then each stimulus card was timed with a stopwatch that was accurate to 1/100 of a second. Time started when the researcher prompted the participant to begin, and time ended when the participant called out the color of the last word item on the card. After the participant completed all the required measures they were debriefed and thanked for their time.

Design and Statistical Analysis
This study employed a 2x2 factorial design, with initial level of depression (initially depressed or initially non-depressed) and vignette condition (enhancing depression or minimizing depression) serving as the two grouping variables. A two-way multivariate analysis of covariance (MANCOVA) was used to examine the effect of initial level of depression and vignette condition on the three dependent variables of BDI-II scores, modified Stroop interference time for neutral words, and modified Stroop interference for depression words. MANCOVA main effects were further analyzed with separate univariate factorial ANCOVAs.
Results

Descriptive statistics generated from the screening procedure for the initially depressed and initially non-depressed participants are presented in Table 1. Frequency information regarding the demographic makeup of both groups, and tests of significance to determine if either of the groups is demographically different, are presented in Table 2. An examination of the demographic characteristics revealed that the depressed and non-depressed participants differed significantly in terms of employment status, χ² (4) = 21.97, p < .01, marital status, χ² (4) = 11.92, p = .02, and level of education, χ² (3) = 7.94, p = .05. Differences between the depressed and non-depressed participants came close to being statistically significant within the category of gender, χ² 2 (1) = 3.70, p = .06. These four variables were controlled for in the MANCOVA.

Examination of Overall MANCOVA

Descriptive data for the dependent variables, categorized by the two independent grouping variables, are presented in Table 3. A significant main effect was observed for initial level of depression (ILD), L = .83, F (3, 62) = 4.20, p = .01, = .17, 1 - b = .84. Additionally, a significant main effect was observed for vignette condition (VC), L = .11, F (3, 62) = 171.10, p < .001, = .89, 1 - b = 1.00. These two main effects warrant additional statistical analysis with separate univariate factorial ANCOVAs for the three dependent variables. Table 4 provides a table of all obtained F-values for the MANCOVA and subsequent post-hoc ANCOVAs.

Examination of Hypotheses

Hypothesis One. The first hypothesis was that depressed participants would have significantly longer Stroop interference times than non-depressed participants on the modified Stroop card with depression related words, regardless of whether they were instructed to respond in a depressed or non-depressed manner. Hypothesis one was tested with a univariate factorial ANCOVA. The ANCOVA showed a significant main effect for ILD on Stroop interference time for depression related words, F (1, 64) = 5.40, p = .02, = .08, 1 - b = .63, which supports the hypothesis that, regardless of which vignette condition participants were exposed to, there is a significant difference in Stroop interference times for depressed participants as compared to non-depressed participants.

Hypothesis Two. The second hypothesis was that depressed and non-depressed participants would not differ from each other on the modified Stroop card with neutral words regardless of whether they were instructed to respond in a depressed or non-depressed manner. This hypothesis was also supported by the statistical evidence, as there was no main effect found for ILD, F (1, 64) = .72, p = .40, = .01, 1 - b = .13, or for VC, F (1, 64) = 1.71, p = .20, = .03, 1 - b = .25. All participants experienced roughly equal amounts of Stroop interference for the neutral word card. Refer back to Table 3 for the mean amount of Stroop interference time for neutral words for each experimental group.

Hypothesis Three. The third hypothesis was that participants who were given the depressed vignette would score significantly higher on the BDI-II than participants who were given the non-depressed vignette, regardless of whether they were initially depressed or non-depressed. Even though a main effect for VC on BDI-II scores was observed, F (1, 64) = 522.19, p < .001, = .89, 1 - b = 1.00, this hypothesis can not be fully supported by ANCOVA data because of a significant ILD x VC interaction, F (1, 64) = 6.69, p = .01, = .10, 1 - b = .72.

This hypothesis was further examined by placing participants’ BDI-II scores into the descriptive categories that their scores reflected. Table 5 shows a frequency count of the descriptive categories grouped by the two independent conditions. When examined by descriptive category, all 18 participants who were initially depressed and told to answer in a non-depressed manner were able to complete the BDI-II showing minimal depression. This same finding is true for all 18 non-depressed participants who were told to respond in a non-depressed manner. Additionally all 18 non-depressed participants, when asked to answer in a depressed manner, scored in the severely depressed range on the BDI-II. Of the 18 depressed participants who were told to respond in a depressed manner, 14 rated themselves as having severe depression, and 4 rated themselves as having moderate depression. This finding lends support to the fourth hypothesis. While initially depressed and non-depressed participants may have slightly different response patterns when fabricating answers, both groups ultimately appear capable of altering their responses to change their current level of depression on the BDI-II. Data analysis indicates partial support for this hypothesis.

Further analysis of this interaction was conducted with traditional t-tests. Within the VC, it appears that depressed participants who were told to respond in a depressed manner (M = 44.28, SD = 14.66) scored significantly higher than depressed participants told to respond in a non-depressed manner (M = 2.94, SD = 3.72), t (34) = 11.60, p < .001. Furthermore,
non-depressed participants told to respond in a depressed manner (M = 53.78, SD = 10.00) scored significantly higher than non-depressed participants told to respond in a non-depressed manner (M = .83, SD = 1.50), t (34) = 22.23, p < .001. Further examination of this interaction was conducted in hypotheses six and seven, which both examine BDI-II scores within ILD.

**Hypothesis Four.** The fourth hypothesis was that, within the initially depressed group, participants who were instructed to respond in a depressed and non-depressed manner would not differ from each other in Stroop response times for either the neutral or depression related word cards. It was expected that the Stroop is resistant at attempts to manipulate one’s responses. This finding is initially supported because there was no observable main effect for VC for both the depressed Stroop word card, F (1, 64) = .66, p = .42, = .01, 1 - b = .13, and the neutral Stroop word card, F (1, 64) = 1.71, p = .20, = .03, 1 - b = .25. Further t-test analysis confirmed that there was no significant difference between depressed Stroop interference times for depressed participants told to respond depressed (M = 17.55, SD = 14.53) as compared to depressed participants told to respond non-depressed (M = 16.16, SD = 13.65), t (34) = .30, p = .77. There was also no significant difference between neutral Stroop interference times for depressed participants told to respond depressed (M = 10.40, SD = 11.23) as compared to depressed participants told to respond non-depressed (M = 8.83, SD = 9.65), t (34) = .45, p = .65.

**Hypothesis Five.** The fifth hypothesis was that, within the initially non-depressed group, participants who were instructed to respond in a depressed and non-depressed manner would not differ from each other in Stroop response times for either the neutral or depression related word cards. The previously mentioned fact that no main effect was observed for VC for both the depressed Stroop word card, F (1, 64) = .66, p = .42, = .01, 1 - b = .13, and the neutral Stroop word card, F (1, 64) = 1.71, p = .20, = .03, 1 - b = .25 also provides support for this hypothesis. Further t-test analysis confirmed that there was no significant difference between depressed Stroop interference times for non-depressed participants told to respond depressed (M = 10.03, SD = 13.85) as compared to non-depressed participants told to respond non-depressed (M = 5.30, SD = 7.22), t (34) = 1.29, p = .21. There was a significant difference between neutral Stroop interference times for non-depressed participants told to respond depressed (M = 9.68, SD = 7.04) as compared to non-depressed participants told to respond non-depressed (M = 5.09, SD = 6.32), t (34) = 2.05, p = .05.

**Hypothesis Six.** The sixth hypothesis was that, within the depressed vignette condition, participants would not differ from each other in BDI-II scores, regardless of whether they were initially depressed or initially non-depressed. This hypothesis is not supported because of the significant ILD × VC interaction, F (1, 64) = 6.69, p = .01, = .10, 1 - b = .72. T-tests were conducted in order to further examine this interaction. It appears that when an initially depressed participant is told to pretend to be depressed, their scores on the BDI-II (M = 44.28, SD = 14.66) are significantly lower than that of an initially non-depressed participant told to pretend to be depressed (M = 53.78, SD = 10.00), t (34) = -2.27, p = .03.

**Hypothesis Seven.** The seventh and final hypothesis was that, within the non-depressed vignette condition, participants would not differ from each other in BDI-II scores, regardless of whether they were initially depressed or initially non-depressed. This hypothesis is also not supported because of the significant ILD × VC interaction, F (1, 64) = 6.69, p = .01, = .10, 1 - b = .72. T-test analysis further revealed that when an initially depressed participant is told to pretend to be non-depressed, their total BDI-II score (M = 2.94, SD = 3.72) is significantly higher than initially non-depressed participants told to pretend to be non-depressed (M = .83, SD = 19.47), t (34) = 2.24, p = .03.

**Discussion**

Study results indicated that the BDI-II was highly sensitive to attempts at manipulation, with varying degrees of vulnerability depending on whether the participant was initially depressed or non-depressed. More importantly, it was demonstrated that while participants were generally successful at manipulating results on the BDI-II, they were unsuccessful at manipulating response patterns on the modified Stroop procedure. Regardless of whether the participant was told to respond in a depressed or non-depressed manner, the modified Stroop was successfully able to differentiate between depressed and non-depressed participants.

It was hypothesized that depressed participants would have significantly longer Stroop interference times for depression related words when compared to non-depressed participants, regardless of whether they were told to respond as depressed or non-depressed. The data analysis supported this hypothesis. This finding was also consistent with previous research that has shown similar findings.25-30
This significant difference suggests that the modified Stroop procedure was able to assess cognitive interference associated with increased levels of depression. This interference was expressed by increased response latency on the depression related word card. Information processing theory suggests that depressed participants exhibited an attention bias for depression related material, and this attention bias was not present in non-depressed participants. The attention bias that depressed participants possessed limited the amount of attention they could use to focus on the color-naming task, thus resulting in increased interference times. In terms of Beck’s model of depression, when depressed participants engaged in the color-naming task, they were also activating their schema for depression by inadvertently reading the word-item. When their schema was activated, it caused processing delays for color naming, thus creating longer interference times when compared to non-depressed participants. Additionally, it has been shown that the modified Stroop procedure was impervious to attempts at manipulating one’s responses. This finding widens the implications discussed by Gottib and McCann26 when they demonstrated that participants would only show increased Stroop interference times if they have a history of depressed mood, and not if their mood is temporarily altered to a state of depression. The current study demonstrated that willful attempts at manipulating responses did not affect Stroop interference times.

It was also hypothesized that depressed and non-depressed participants would not differ from each other on the modified Stroop procedure with neutral words, regardless of whether they were told to respond as depressed or non-depressed. This hypothesis was supported. This finding was consistent with research on the modified Stroop procedure that demonstrated participants showed equal amounts of interference for neutral material, regardless of whether or not the participant was depressed.25-27, 29,30

In terms of information processing theory, depressed and non-depressed participants did not show any attention biases towards neutral material. Therefore, they were all able to focus the same amount of attention on the color-naming task. Beck’s model of depression would assert that these neutral words did not activate the participants’ schema for depression. Therefore, everyone experienced roughly the same amount of interference because depressed participants did not have the handicap of additional schema activation when compared to non-depressed participants. The task was equally challenging to all participants.

It was hypothesized that, within the depressed and non-depressed groups, participants would not show a difference on Stroop interference times, regardless of whether they were instructed to respond in a depressed manner or non-depressed manner. This finding was supported and provided evidence for the ability of the modified Stroop procedure to assess whether a participant was depressed or non-depressed, even if they were trying to mask their true mood. Thus, depressed and non-depressed participants were equally poor at influencing their response times on the modified Stroop procedure and this was consistent with prior research that examines the modified Stroop procedure’s ability to measure state versus trait depression.34 If the modified Stroop procedure assessed trait depression, as discussed by William and Nulty34, then not only would Stroop interference times be stable over an extended period of time, but they would also be resistant to attempts at manipulation. The modified Stroop procedure indirectly assessed the development of a person’s schema for depression. As such, it is difficult to alter over time, and difficult for a person to willfully manipulate. There was one significant difference within the t-test analysis examining neutral Stroop times. It appears that there is a statistically significant difference between neutral Stroop interference times for non-depressed participants told to respond depressed as compared to non-depressed participants told to respond non-depressed. While this finding is statistically significant, the difference between the two groups is only 4.59 seconds. This difference is slightly less than the difference between depressed participants told to respond depressed as compared to depressed participants told to respond non-depressed (4.73 seconds). This significant difference may not be due to any real difference in performance, and may be due to random statistical effects.

These findings demonstrate that the modified Stroop procedure could be a useful tool for depression assessment, especially if a researcher or clinician suspects the participant may be engaging in deception. From a practical standpoint, all participants (whether depressed or non-depressed) possess equally weak abilities in creating deception on the modified Stroop procedure.

According to information processing theory, this suggests that depressed participants had a stable information bias for depression. This bias was present regardless of whether or not a person was told to “act” depressed or non-depressed. It did not matter how the participant willingly tried to act, their attention bias created interference without the depressed participant being able to control this effort.
Finally, it was hypothesized that participants who were told to respond in a depressed manner would have significantly higher scores on the BDI-II than participants who were told to respond in a non-depressed manner, regardless of whether or not they were initially depressed or initially non-depressed. Analysis of the data provided mixed support for this hypothesis. If one examines actual scores on the BDI-II, it appears that depressed participants scored slightly higher than non-depressed participants when told to respond in a depressed manner. Support for this hypothesis occurs when one examines the descriptive categories that these BDI-II scores represent. Ultimately, when a depressed participant was asked to respond in a non-depressed manner, their scores fell in the non-depressed range. When a depressed participant was asked to respond in a depressed manner, all scores fell in the depressed range. This shows that the BDI-II was highly sensitive toward attempts at manipulation, and was consistent with past research specifically examining a person’s ability to fabricate results on a BDI.12,13 This finding also confirms suspicions that were raised in previous research regarding a person’s ability to fabricate answers on paper and pencil psychometric tests.12-15 In the present study, all participants were able to change their response pattern at will. The fact that non-depressed participants were able to successfully fabricate responses on the BDI-II supported the research of Lees-Haley13 when it was observed that non-depressed participants were adept at altering their responses on the BDI in order to appear depressed. The current study expands these findings one step further by showing that individuals with high levels of depression were equally adept at appearing non-depressed on the BDI-II. These results show that participants of varying levels of education and various ethnicities were quite capable of deliberately manipulating their responses on the BDI-II. Results relevant to the sixth and seventh hypotheses revealed some unexpected interaction findings when initially depressed and non-depressed participants were asked to fake responses in a particular direction on the BDI-II. However, the overall implication of these results is to further support the vulnerability of the BDI-II to intentional response alteration.

Limitations of the Current Study
One potential weakness of the study was that a categorical approach was utilized to distinguish between depressed and non-depressed participants. Most current research shows that depression is a continuous variable, with people falling somewhere between two extremes of either “completely depressed” or “completely non-depressed.”46,47 The choice to examine depression categorically was made because this reflects the current state of clinical practice in psychology. In the clinical sphere, individuals are usually diagnosed categorically, not along a continuum. Another potential limitation is our use of the SDS and the RHRSD self-report inventories in the selection of the depressed sample. It might be argued that if we are critiquing the BDI-II as too obtrusive, why use similar measures to select depressed participants? This is a valid concern and yet it is hoped that by relying on two inventories rather than one, the selection error would be somewhat lessened. The lack of available alternative methods of assessing depression, to some extent, forced a reliance on the type of measures that we critique.

An additional limitation is that this study did not include a control group that was not asked to fake responses on either the BDI-II or the Stroop. Including such a control group would be important in future research as it would provide baseline data from which one could assess the degree to which responses change on the BDI-II and Stroop with willful efforts to distort. Our expectation, based upon current findings, is that the magnitude of change of depression level would be greater on the BDI-II than on the Stroop, but this is an important empirical issue that awaits further research.

Clinical Implications and Suggestions for Research
The results of this study provide evidence to practicing clinicians that they may wish to consider the use of alternative measures to the BDI-II for assessing depression when the client they are examining may be motivated to distort their responses. Since the modified Stroop procedure appears to be relatively impervious to attempts at manipulating response patterns, it can be considered as a possible aid in guiding a clinician toward a diagnosis. For acceptance as a clinical measure of depression, it is important to understand what the BDI-II and the modified Stroop procedure measure, and how these scales are used. A needed, but involved, avenue of research would be further validation of the modified Stroop procedure as a depression assessment tool, and perhaps the development of standardized Stroop interference times, which would generate cutoff scores that would indicate whether or not a person is depressed and their level of depression. The establishment of cutoff scores for various levels of depression is a complex undertaking. One possible way of establishing cutoff scores is to have various depressed, non-depressed, samples take a series of depression assessment devices along with the Stroop, and to also include samples with forms of psychopathology other than depression. One could then compare various levels of
depression, across multiple samples, with response latencies on the Stroop. One would then be in a defensible position to establish the Stroop’s incremental and discriminant validity after having assessed a large number of varied samples. Due to the lack of standardized items and standardized presentation protocols for the modified Stroop procedure, the amount of Stroop interference can vary greatly between studies. Research in this area would have to focus on identifying universally accepted depression related words, and developing average Stroop interference times based on demographic characteristics for both depressed and non-depressed populations. Eventually, one may be able to determine if a person exhibits symptoms of depression by examining an individual’s discrepancy between modified Stroop interference times for neutral words and modified Stroop interference times for depression related words.

A potential methodological improvement might be to present individual stimuli by computer rather than having participants respond to stimuli on cards, as was done in the current study. Jensen,48 for example, presents research on visual and auditory reaction times for accessing various cognitive skills and ability areas. It should be noted that although the Stroop is comparatively impervious to willful attempts to alter responses, Jensen’s work suggest that some degree of response alteration is possible even on “reaction time” tasks such as the Stroop. A large amount of research using the Stroop to assess emotional states, however, has not relied on computer presentations and it is for this reason that the card format presentation was used here. Nevertheless, the ability to measure cognitive task performance within milliseconds makes the use of computers a promising avenue of future research.

Conclusion

In conclusion, the modified Stroop procedure proved to be resistant to participants’ attempts at fabricating responses in order to appear depressed or non-depressed. The BDI-II was susceptible to attempts at manipulation. While depressed and non-depressed participants fabricate responses on the BDI-II with slightly different response patterns, all participants are capable of achieving the ultimate goal or appearing depressed or non-depressed. The findings of this study provide empirical evidence that the validity of the BDI-II is weakened in specific situations, such as when a participant is motivated to fabricate responses, and there are alternative measures like the modified Stroop procedure that still successfully differentiate between depressed and non-depressed participants in these situations.

References

4. Carro IL, Bernal IL, Vea HB: La Depresion en Cuba: Validacion del Inventario de Depresion de Beck y de la Escala de Actitudes Disfuncionales en poblacion Cubana [Depression in Cuba: Validation of Beck Depression Inventory (BDI) and the Dysfunctional Attitudes Scale (DAS-A) with Cuban population]. Avances en Psicologia Clinica Latinoamerica. 1998;16:111-120.
13. Lees-Haley PR: Malingering traumatic mental


Illustrations

Illustration 1

Table 1. Age and Initial Screening Measure Data on the SDS and RHRSD (N = 72)

<table>
<thead>
<tr>
<th>Screening Variable</th>
<th>Initial Level of Depression</th>
<th>Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Non-Depressed</td>
<td>Depressed</td>
</tr>
<tr>
<td>Age</td>
<td>42.72</td>
<td>46.36</td>
</tr>
<tr>
<td>SDS Score</td>
<td>28.89</td>
<td>50.06</td>
</tr>
<tr>
<td>RHRSD Score</td>
<td>6.00</td>
<td>26.81</td>
</tr>
</tbody>
</table>

Note.  
M = Mean  
SD = Standard Deviation  
SDS = Zung Self-Rating Depression Scale  
RHRSD = Revised Hamilton Rating Scale for Depression  
a_n = 36 for each initial mood condition.

Scores on the SDS (Zung, 1965) are divided into four categories: 20 to 39 is non-depressed, 40 to 47 is mildly depressed, 48 to 55 is moderately depressed, and 56 and above is severely depressed. Scores on the RHRSD (Warren, 1994) are divided into four categories: 0 to 10 is non-depressed, 11 to 16 is minor depression, 17 to 25 is major depression, and 26 and above is severe depression.
Illustration 2

Table 2. Frequency Data and Tests of Significance for Non-Depressed and Depressed Participants (N = 72)

<table>
<thead>
<tr>
<th>Demographic Variable</th>
<th>Initial Level of Depression(^a)</th>
<th>Statistics</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Non-Depressed</td>
<td>Depressed</td>
<td>df</td>
<td>(\chi^2)</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td>1</td>
<td>3.70</td>
</tr>
<tr>
<td>Male</td>
<td>19</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>17</td>
<td>26</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
<td>4</td>
<td>5.54</td>
</tr>
<tr>
<td>Asian</td>
<td>1</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>2</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caucasian</td>
<td>30</td>
<td>23</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>2</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
<td>3</td>
<td>7.94</td>
</tr>
<tr>
<td>High School/GED</td>
<td>9</td>
<td>19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Some College</td>
<td>8</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>College Diploma</td>
<td>10</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Graduate School</td>
<td>9</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Marital Status</td>
<td></td>
<td></td>
<td>4</td>
<td>11.92</td>
</tr>
<tr>
<td>Single</td>
<td>10</td>
<td>19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>23</td>
<td>9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Divorced</td>
<td>0</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Separated</td>
<td>3</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Widowed</td>
<td>0</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employment Status</td>
<td></td>
<td></td>
<td>4</td>
<td>21.97</td>
</tr>
<tr>
<td>Work Full Time</td>
<td>21</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Work Part Time</td>
<td>6</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unemployed</td>
<td>5</td>
<td>10</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. \(^a\) \(n = 36\) for each initial mood condition.
## Illustration 3

Table 3. Descriptive Data for the BDI-II and the Modified Stroop Interference Times Categorized by Initial Level of Depression and Vignette Condition (N = 72)

<table>
<thead>
<tr>
<th>Vignette Condition</th>
<th>Initial Level of Depression$^a$</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Non-Depressed</td>
<td>Depressed</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$M$</td>
<td>$SD$</td>
<td>$M$</td>
</tr>
<tr>
<td>Non-Depressed Vignette</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BDI-II Score</td>
<td>.83</td>
<td>1.47</td>
<td>2.94</td>
</tr>
<tr>
<td>Neutral Stroop Interference</td>
<td>5.08</td>
<td>6.32</td>
<td>8.83</td>
</tr>
<tr>
<td>Depressed Stroop Interference</td>
<td>5.30</td>
<td>7.22</td>
<td>16.16</td>
</tr>
<tr>
<td>Depressed Vignette</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BDI-II Score</td>
<td>53.78</td>
<td>10.00</td>
<td>44.28</td>
</tr>
<tr>
<td>Neutral Stroop Interference</td>
<td>9.68</td>
<td>7.04</td>
<td>10.41</td>
</tr>
<tr>
<td>Depressed Stroop Interference</td>
<td>10.03</td>
<td>13.85</td>
<td>17.55</td>
</tr>
</tbody>
</table>

*Note.* $M =$ Mean  
$SD =$ Standard Deviation  
BDI-II = Beck Depression Inventory – Second Edition  
$^a$ = 18 for each table cell.  
Scores on the BDI-II (Beck et al., 1996) are divided into four descriptive categories: 0 to 13 is minimal depression, 14 to 19 is mild depression, 20 to 28 is moderate depression, and 29 to 63 is severe depression.
Illustration 4

Table 4. MANCOVA and Post-Hoc Univariate ANCOVA Tests for Initial Level of Depression and Vignette Condition While Controlling for Gender, Level of Education, Marital Status, and Employment Status (N = 72)

<table>
<thead>
<tr>
<th></th>
<th>df</th>
<th>$\Lambda$</th>
<th>$F$</th>
<th>$p$</th>
<th>$\eta_p^2$</th>
<th>$1-\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Overall MANCOVA</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Initial Level of Depression (ILD)</td>
<td>3, 62</td>
<td>.83</td>
<td>4.20</td>
<td>.01</td>
<td>.17</td>
<td>.84</td>
</tr>
<tr>
<td>Vignette Condition (VC)</td>
<td>3, 62</td>
<td>.11</td>
<td>171.10</td>
<td>&lt;.001</td>
<td>.89</td>
<td>1.00</td>
</tr>
<tr>
<td>ILD vs VC</td>
<td>3, 62</td>
<td>.90</td>
<td>2.31</td>
<td>.09</td>
<td>.10</td>
<td>.55</td>
</tr>
<tr>
<td><strong>Depressed Word Stroop Interference</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ILD</td>
<td>1, 64</td>
<td>—</td>
<td>5.40</td>
<td>.02</td>
<td>.08</td>
<td>.63</td>
</tr>
<tr>
<td>VC</td>
<td>1, 64</td>
<td>—</td>
<td>.66</td>
<td>.42</td>
<td>.01</td>
<td>.13</td>
</tr>
<tr>
<td>ILD vs VC</td>
<td>1, 64</td>
<td>—</td>
<td>.63</td>
<td>.43</td>
<td>.01</td>
<td>.12</td>
</tr>
<tr>
<td><strong>Neutral Word Stroop Interference</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ILD</td>
<td>1, 64</td>
<td>—</td>
<td>.72</td>
<td>.40</td>
<td>.01</td>
<td>.13</td>
</tr>
<tr>
<td>VC</td>
<td>1, 64</td>
<td>—</td>
<td>1.71</td>
<td>.20</td>
<td>.03</td>
<td>.25</td>
</tr>
<tr>
<td>ILD vs VC</td>
<td>1, 64</td>
<td>—</td>
<td>.83</td>
<td>.37</td>
<td>.01</td>
<td>.15</td>
</tr>
<tr>
<td><strong>BDI-II Scores</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ILD</td>
<td>1, 64</td>
<td>—</td>
<td>1.68</td>
<td>.20</td>
<td>.03</td>
<td>.25</td>
</tr>
<tr>
<td>VC</td>
<td>1, 64</td>
<td>—</td>
<td>522.19</td>
<td>&lt;.001</td>
<td>.89</td>
<td>1.00</td>
</tr>
<tr>
<td>ILD vs VC</td>
<td>1, 64</td>
<td>—</td>
<td>6.69</td>
<td>.01</td>
<td>.10</td>
<td>.72</td>
</tr>
</tbody>
</table>
Illustration 5

Table 5. Frequency Data for BDI-II Descriptive Categories Grouped by Initial Level of Depression and Vignette Condition (N = 72)

<table>
<thead>
<tr>
<th>Vignette Condition</th>
<th>Initial Level of Depression&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Non-Depressed</td>
</tr>
<tr>
<td>Non-Depressed Vignette</td>
<td></td>
</tr>
<tr>
<td>Minimal Depression</td>
<td>18</td>
</tr>
<tr>
<td>Mild Depression</td>
<td>0</td>
</tr>
<tr>
<td>Moderate Depression</td>
<td>0</td>
</tr>
<tr>
<td>Severe Depression</td>
<td>0</td>
</tr>
<tr>
<td>Depressed Vignette</td>
<td></td>
</tr>
<tr>
<td>Minimal Depression</td>
<td>0</td>
</tr>
<tr>
<td>Mild Depression</td>
<td>0</td>
</tr>
<tr>
<td>Moderate Depression</td>
<td>0</td>
</tr>
<tr>
<td>Severe Depression</td>
<td>18</td>
</tr>
</tbody>
</table>

<sup>a</sup> n = 18 for each table cell.

Scores on the BDI-II (Beck et al., 1996) are divided into four descriptive categories: 0 to 13 is minimal depression, 14 to 19 is mild depression, 20 to 28 is moderate depression, and 29 to 63 is severe depression.
Illustration 6

Appendix A.Vignette that Minimizes Depressive Symptoms

In this study, you are being asked to imagine that you are having a psychological evaluation by your
insurance company. Your insurance has instituted a new policy, and has mandated everyone to
undergo a psychological evaluation to determine if any of the company’s customers require
psychotherapy. One way a person can receive therapy is if they have an emotional disorder, such as
depression. However, you feel that you do not need any therapy and are able to handle your life
problems on your own. The psychologist conducting the evaluation has presented you with two
tests and tells you that these tests will determine whether or not you have an emotional disorder.

**Complete the following tests so the results ensure that it will appear that you do not have an emotional disorder.**
Illustration 7

Appendix B. Vignette that Maximizes Depressive Symptoms

In this study, you are being asked to imagine that you are having a psychological evaluation by your insurance company. Your insurance has instituted a new policy, and has mandated everyone to undergo a psychological evaluation to determine if any of the company’s customers require psychotherapy. Coincidently, you feel that you need therapy because you are unable to handle your life problems on your own. The psychologist conducting the evaluation has presented you with two tests and tells you that these tests will determine whether or not you have an emotional disorder. Complete the following tests so the results ensure that it will appear that you have an emotional disorder.
Disclaimer

This article has been downloaded from WebmedCentral. With our unique author driven post publication peer review, contents posted on this web portal do not undergo any prepublication peer or editorial review. It is completely the responsibility of the authors to ensure not only scientific and ethical standards of the manuscript but also its grammatical accuracy. Authors must ensure that they obtain all the necessary permissions before submitting any information that requires obtaining a consent or approval from a third party. Authors should also ensure not to submit any information which they do not have the copyright of or of which they have transferred the copyrights to a third party.

Contents on WebmedCentral are purely for biomedical researchers and scientists. They are not meant to cater to the needs of an individual patient. The web portal or any content(s) therein is neither designed to support, nor replace, the relationship that exists between a patient/site visitor and his/her physician. Your use of the WebmedCentral site and its contents is entirely at your own risk. We do not take any responsibility for any harm that you may suffer or inflict on a third person by following the contents of this website.