

2004

Effects of Intervening Items on Self-Esteem Tests: A Study of Face Validity

Douglas Wayne Snuffer II

Follow this and additional works at: <http://mds.marshall.edu/etd>



Part of the [Personality and Social Contexts Commons](#)

Recommended Citation

Snuffer, Douglas Wayne II, "Effects of Intervening Items on Self-Esteem Tests: A Study of Face Validity" (2004). *Theses, Dissertations and Capstones*. Paper 882.

Effects of Intervening Items on Self-Esteem Tests: A Study of Face Validity

**Thesis submitted to
The Graduate College of
Marshall University**

**In partial fulfillment of the
Requirements for the degree of
Master of Arts
Psychology**

by

Douglas Wayne Snuffer II

**Dr. David Trumpower, Committee Chairperson
Dr. Steven Mewaldt
Dr. Christopher LeGrow**

Marshall University

Fall 2004

This thesis was accepted on November 23 2004
Month Day Year

as meeting the research requirements for the master's degree

Dr. David Trumpower, Committee Chairperson

Dr. Steven Mewaldt, Committee Member

Dr. Christopher LeGrow, Committee Member

ABSTRACT

Effects of Intervening Items on Self-Esteem Tests: A Study of Face Validity

by Douglas Wayne Snuffer II

The domain of self-esteem testing has often been thought to suffer from problems concerning demand characteristics. High face validity of many common tests of self-esteem, such as the Rosenberg (1965) Self-Esteem Scale, may influence responding of test takers, such that they respond in a socially desirable way. The current study investigated the effects of inserting “neutral” intervening items between items of the Self-Esteem Scale, creating a test that is less face valid. No significant effect was discovered concerning the primary hypothesis, though other significant results were found. Explanations and implications of the obtained results are discussed.

DEDICATION

The author wishes to dedicate this work to anyone who said that he was too lazy to ever accomplish anything. Particularly, his father.

ACKNOWLEDGEMENTS

I would like to thank the following people for their direct assistance in the completion of this thesis. First of all, thanks to Christina Johnson, the love of my life, the best woman I know, my favorite person in the world, for letting me use her computer, among other things. Dr. Dave Trumpower, for his relentless pursuit of statistical perfection, and for making me think in ways I had never known. Dr. Steve Mewaldt and Dr. Chris LeGrow for sitting on my committee, and giving helpful suggestions along the way. Dr. Pam Mulder, for her assistance in my pilot study. Leasha Trimble, for her assistance in test construction. Last but not least, I thank myself, for writing this god forsaken thing.

I would also like to thank the following people for shaping me into the scholar that I have become. Dr. Carl Chapman, Dr. Fred Bailey, Dr. Roger Shepperd, Mary Beckett, Dr. Debbie Vause, Dr. Rachel Ivey, Peggy Baldwin, Ken Scott, Don Davis, Dr. Irene Moser, Dr. Liz Roth, Jack Deskins, Dr. Angelique Edwards, Dr. Jim White, Jan Westerik, Dr. George Towers, Dr. Roberta Salyers, Dr. Dannette Light, Dr. Elizabeth Schiltz, Dr. Kristen Jenkins, Lisa Karnes, Dr. Kathleen Connor (R.I.P), Mike McDaniel, Dr. Steve Salinger (R.I.P.), Bill Hartling, Dr. Jan Krell, Dr. Rod Klein, Dr. Karen Griffee, Dr. Joe Wyatt, Dr. Helen Linkey, Birgit Shanholtzer, and any other college faculty I've met along the way whose name I have forgotten. Special thanks to Dr. Wilber Jones (Retired) of the Concord College chemistry department, for showing me that psychology was the best route for me.

Finally, a thank you to my daughter Alexis, for keeping me alive when I had no other reason.

TABLE OF CONTENTS

ABSTRACT.....	iii
DEDICATION.....	iv
ACKNOWLEDGEMENTS.....	v
TABLE OF CONTENTS.....	vi
LIST OF TABLES.....	vii
CHAPTER I – Introduction and Review of Literature.....	1
CHAPTER II – Methods.....	16
Participants.....	16
Materials.....	16
Procedure.....	18
CHAPTER III – Results.....	19
CHAPTER IV – Discussion.....	28
References.....	35
Appendices.....	39

LIST OF TABLES

Table 1. - <i>Pairwise Comparisons for Ratings of Honesty as Perceived by Others</i>	22
Table 2. - <i>Paired Samples t-tests for self vs. others Honesty Ratings</i>	23
Table 3. - <i>Paired Samples t-tests for self vs. others Self-Esteem Ratings</i>	24
Table 4. - <i>Correlations Between SES Scores, SELF-SE, and SE-OTHERS</i>	24
Table 5. - <i>Pairwise Comparisons for Self-Esteem Difference Scores</i>	26
Table 6. - <i>Frequencies of Identifying Self-Esteem as a Recurring Theme by Group</i>	27

CHAPTER I

Introduction and Review of Literature

Effects of Intervening Items on Self-Esteem Tests: A Study of Face Validity

One of the more researched aspects of personality over the past century has been self-esteem. The scientific study of self-esteem appears to have its roots in the writings of William James, who may have coined the term. In his text, *Principles of Psychology* (1890), James lays out a formula that both defines and calculates self-esteem.

$$\text{self - esteem} = \frac{\text{success}}{\text{pretensions}} \quad (1)$$

In this way, James defines self-esteem as a function of our successes in comparison to our pretensions, or aspirations. In other words, if a person's level of success is similar to the success that the person wishes for, then that person has high self-esteem. In order to increase self-esteem, one would be required to either raise their level of success, or lower their standards.

James refers to his paradigm for self-esteem as a "ratio" or "fraction" (p. 310), implying that some numerical value can be used when discussing a person's self-esteem. Thus began the mission of the psychologist to quantify the construct. A number of problems no doubt arise however when this mission is tackled. First of all, it is of utmost importance to first operationally define success and pretensions in order to quantify those values from which James' formula derives its basis. Secondly, what if self-esteem need not be concerned with success and pretension, or perhaps more accurately, needs to be concerned with several other concepts as well? How do we measure these other concepts?

Upon consideration of these issues, it soon becomes obvious that the true problem

in measuring self-esteem is that it first must be well-defined. It is not enough to take James' writings and establish definitions of self-esteem from them merely because James stated it, although this was the case for many years (Pelham, 1995). In order for a concept to be successfully defined, it must be supported by other researchers and their work (Ward, 1996). Establishing a solid definition of self-esteem, then, is by no means a simple task.

To add difficulty to the work ahead, James himself offered a second view of self-esteem, one that had nothing to do with pretensions or successes. In this view, James referred to self-esteem as an "average tone of self-feeling" (p. 306), which greater reflects the consensus today about what self-esteem is. James' second view of self-esteem relates to what we now identify as *global self-esteem*, whereas his first view, involving success and pretensions, is typically thought of as *domain-specific self-esteem*.

Part of the problem in defining self-esteem may result from the different approaches used to define it (Reasoner, 2004). That is to say, different researchers have defined self-esteem in terms of different things. For example, behaviorists may define self-esteem in terms of dominant behaviors (though traditionally, behaviorists had no concern with things like self-esteem, due to its unobservable nature [Ward, 1996]), while humanists may define self-esteem in terms of actual versus ideal self, much like William James. Others (e.g., Rosenberg, 1965) define self-esteem simply in terms of positive and negative attitudes. Differing viewpoints such as these have made it difficult to create a consensus definition. Because of this, the measurement of self-esteem has taken on a variety of forms. Most often however, self-esteem is tested with a self-report questionnaire, in which respondents state how much a given statement applies to them

(Hoyle, Kernis, Leary, & Baldwin, 1999).

Tests of self-esteem, though not specifically stated to be self-esteem tests, were first developed by Abraham Maslow in the late 1930's/early 1940's. Maslow originally conceptualized self-esteem in humans as a feeling of dominance, closely related to his observations of self-confidence in dominant monkeys (1937, as cited in Cullen, 1997). A few years later (1940), Maslow developed a test of *dominance-feeling* (self-esteem), which seemed to correlate well with sex drive in women. This established an anchor connecting dominance to self-esteem, as dominant monkeys also tend to have active sex lives (Cullen, 1997).

Maslow's test of dominance-feeling consists of two sections. The first section, comprised of 27 multiple-choice questions, asks test takers (who were exclusively women for Maslow's original 1940 study) how they feel about certain issues, or how frequently they experience a particular feeling. Examples include: "What is your attitude toward men who *look* as if they could be brutal?" and "How frequently are you embarrassed?" Answers were weighted differently for each question. That is, the possible responses for one question could have values of 3, 5, 2, -3, or -2; the possible responses for the next question could have values of -3, 0, 2, 6, or 1. These weightings were derived from tables found in Dr. John Flanagan's (1935, as cited by Maslow, 1940) work on factor analysis in personality. The test taker is given the value that corresponds with their response to that particular question. Values are then summed, with higher scores being indicative of higher feelings of dominance.

The second section of the test consists of 25 items in which the test taker is to declare how much they like or dislike something. A five-point Likert scale is used,

ranging from -2 to +2, with -2 implying strong dislike, and +2 implying strong liking. Once again, answers are weighted differently for each question. A response of -2 for one question could garner a score of 3, while the same response for another question could be valued as -6. Once again, higher scores indicated higher feelings of dominance.

As groundbreaking as Maslow's test was, it is not without flaw, and Maslow noticed and acknowledged this fact. Maslow cautions readers in the use of his test, stating that the test scores do not always reflect personal interviews, and thus should not necessarily be used with participants who cannot be interviewed. He also writes, "...this test should not be used as a substitute for the most valid way of studying personality, namely, to study the personality directly, as a whole" (1940, p.264).

In the 1950's, much research was done on the effects of self-esteem on a person's life. Self-esteem and self-concept were shown to be related to ethnocentrism (Pearl, 1954), level of social interaction (Manis, 1955), and delinquency (Reckless, Dinitz, & Kay, 1957), as well as many other important life facets. While no serious advancements in the measurement of self-esteem were made, it became evident that self-esteem was now an important issue in psychological research. Thus, the urgency of creating a better measurement device became even greater.

In the 1960's, two important books on self-esteem were written, and with them came two new tests of self-esteem, both of which are widely used today. Rosenberg's *Society and the Adolescent Self-Image* (1965), and Coopersmith's *The Antecedents of Self-Esteem* (1967) both sought to find the influences and consequences of self-esteem, largely in youth populations. Both Rosenberg and Coopersmith define self-esteem as an attitude, largely dealing with feelings of worthiness. Due to this definition, the Rosenberg

Self-Esteem Scale (SES) and the Coopersmith Self-Esteem Inventory (SEI) both measure self-esteem by assessing a person's attitude about themselves. This task is accomplished by presenting a respondent with a list of statements (e.g., I think that I am a good person), then asking the respondent to what extent the statement applies to them. The SES consists of 10 items of this nature, while the SEI (adult form) is comprised of 25 items.

Currently, theory and research have returned to ideas of old about self-esteem, relating to the work of Maslow, and even William James. One of the more recent theories about the nature of self-esteem is the sociometer theory (Leary, Tambor, Terdal, & Downs, 1995). One of the postulates of the sociometer theory is that the self monitors the social environment for cues that indicate social rejection. Lowered self-esteem can be the result. Self-esteem, then, is the sociometer which gauges success in gaining social acceptance. Sociometer theory is similar to Maslow's dominance theory in that both of these approaches regard self-esteem as a monitor for an individual's interpersonal relationships (Leary, Cottrell, & Phillips, 2001).

Recent research has also supported the theories of James, presupposing that self-esteem is related to a person's ratio of success and pretensions. Morretti and Higgins (1990) showed that the number of differences that exist between the actual self (i.e., successes) and the ideal self (i.e., pretensions) is negatively related to self-esteem. Pelham (1995) found that being relatively successful in areas thought to be important by the individual, which he termed *differential importance*, was related to higher levels of self-esteem. Pelham's findings support an interesting point that was originally proposed by James (1890). According to James, failure in areas that one cares about will no doubt have an impact on that person's self-esteem. In other areas, failure will have less of an

effect. For example, if one were aspiring to be a great baseball player, but lacked the coordination to hit the ball with the bat consistently, self-esteem would suffer. If the same person failed a piloting test, self-esteem would not suffer as much, since that person wants to be a baseball player, not a pilot.

Why Measure Self-Esteem?

While some people (e.g., behaviorists) may argue that there is little point in measuring self-esteem due to its unobservable nature, there should be very little argument that self-esteem is an important aspect of a person's life. Numerous studies have linked low self-esteem, as measured by the SES (e.g., Silverstone & Salsali, 2003; Rosenberg, 1965) and by the Unconditional Self-Regard Scale (Smith & Betz, 2002), with depression. With depression being so highly related to suicide (Harris & Barraclough, 1997, as cited in Goldney, 2003), it becomes obvious that an accurate measurement of self-esteem is needed. With accurate measurements of self-esteem, clinicians can better identify those at risk of suicide, and provide intervention accordingly.

An additional use of self-esteem testing could be in relationship counseling. Self-esteem, as measured by the SES, has been negatively linked to manic love ($r = -.25$) and positively linked to erotic love ($r = .15$; Campbell, Foster, & Finkel, 2002). Self-esteem, as measured by a two-item scale found to moderately correlate ($r = .54$) with the SES, is also positively correlated with relationship satisfaction in men (Hendrick, Hendrick, & Adler, 1988). While it would be pointless to test self-esteem to discover a problem in a relationship (going to counseling should, by itself, provide enough evidence that a problem exists), testing could be quite useful in discovering progress. That is, if self-esteem rises during the course of counseling, it could be indicative of less

relationship problems, possibly resulting from the counseling provided.

The Coopersmith Self-Esteem Inventories

Originally intended for young children, the SEI child form is a 58-item measure; fifty of the items are used to assess self-esteem, while the other eight items are used as a lie scale. Test takers are asked to read each of the statements, then place a checkmark in a blank either labeled as “*Like me*” or “*Unlike me*” (Coopersmith, 1967). Logically, responding with “*Like me*” to statements such as “I’m pretty sure of myself” (p. 265) results in higher scores of self-esteem. Some items on the SEI, however, are negatively worded and thus, are reverse scored. Each item is worth 2 points, making possible total scores range from 0 - 100, with higher scores relating to higher self-esteem. The lie scale, which is used as a measure of defensiveness, is comprised of items that are worded *all-case inclusively*, such as “I never get scolded” and “I always tell the truth” (p. 266). Test takers who respond to a majority of these items with “*Like me*” are thought to be defensive, and their self-esteem scores may be less meaningful.

The adult form of the SEI is, ironically, shorter than its child counterpart and has no lie scale. It consists of 25 items which are similar in nature to the items of the child scale. The response format is identical. Each item is worth 4 points, allowing scores to range from 0 - 100. Again, higher scores relate to higher self-esteem.

Both forms of the SEI are easily administered and scored. Reliability for the SEI is rather high, with *K-R 20* ranging from .87 to .92 (Sewell, 1985). However, the adult form of the SEI suffers from several problems. The most important of these is the notion that self-esteem measures may be confounded by *social desirability*, the tendency to respond according to social appropriateness (Peterson & Austin, 1985). It is possible then

that SEI test takers may try to “fake good” to give the impression of having high self-esteem. Empirical support for this suspicion is provided by Francis (1997), who found that SEI scores were significantly positively correlated ($r = .08$) with the Lie scale of the Junior Eysenck Personality Inventory. This correlation, while significant, is rather low, which may imply that some people fake good when responding to questions about their self-esteem, but not most people.

Additionally, scores on the SEI have been found to positively correlate ($r = .37$) with extraversion (Francis, 1997), implying that low self-esteem is a characteristic of introverts. While this implication may be true, it is also possible that the SEI is biased toward extraverts. Therefore, introverts who respond to the SEI could be misinterpreted as having low self-esteem. Consistent with the idea that self-esteem and extraversion are related, Rosenberg (1965) found that adolescents with low self-esteem were less likely to participate in extracurricular activities.

The Rosenberg Self-Esteem Scale

The SES is a 10-item self-report questionnaire designed to measure a person’s self-esteem. It was created as part of a multi-questionnaire battery, originally administered to 5,024 high school students in New York State. The SES was originally intended to measure self-esteem in adolescents. Rosenberg’s reasons for using adolescents for his original sample are multi-fold, one of which being that adolescents tend to focus on their self-image more than others. Today, the SES is used in numerous settings, from schools to clinics, and with people of various ages, from adolescence into adulthood (Measure: Rosenberg self-esteem scale [SES], n.d.).

The design of the SES is that of a survey. Each of the 10 items provides a

statement such as, “On the whole, I am satisfied with myself” and “I feel I do not have much to be proud of.” Possible responses come in the form of a 4-point Likert scale: *strongly agree* (SA), *agree* (A), *disagree* (D), and *strongly disagree* (SD). Test takers respond by reading the given statement and circling the response that best corresponds with their level of agreement with the statement. The 10 items of the SES are listed in Appendix A.

Scoring the SES involves giving a numerical value to each of the responses, then adding the values across the 10 items. Typically these values are: strongly agree = 4, agree = 3, disagree = 2, and strongly disagree = 1. Five of the items on the SES are negatively worded (i.e. agreement implies lower self-esteem). These five items are reverse scored. Therefore, final scores can range from 10 – 40, with higher scores indicating higher self-esteem.

Pros. One of the many positive points about the SES is its utilization of a 4-point Likert scale. Typically, the use of a 4-point scale eliminates neutral response bias. Additionally, having a 4-point scale, as opposed to the 2-point “like me/unlike me” scale of the SEI, increases the per-item sensitivity of the test.

Also, the inter-item consistency of the SES is exceptionally high for such a short test. Cronbach alphas for the SES have been shown to range from .84 (Gudjonsson & Sigurdsson, 2003) to .90 (Watson, Suls, & Haig, 2002). In addition, test-retest reliability for the SES was also found to be high, ranging from .82 to .85 (Fleming & Courtney, 1984).

Another attractive quality of the SES is its short length. Being only 10 items long, the SES is not burdened with the problem of testing fatigue. Testing with the SES

typically requires five minutes or less. Scoring and interpretation are also quickly accomplished.

Convergent validity of the SES has been established in previous studies. SES scores were found to significantly correlate with depression ($r = -.54$) and suicidal ideation ($r = -.50$; de Man & Gutierrez, 2002). Additionally, Gudjonsson and Sigurdsson (2003) found that self-esteem scores, as measured by the SES, correlated significantly ($r = -.41$) with compliance. Given that people with low self-esteem feel that their wants or desires are less important than those of others, this correlation is logical.

Cons. It is possible that the length of the test is problematic for the SES as well. While the per-item sensitivity of the SES may be greater than that of the SEI, the SEI allows for nearly as many possible scores as the SES, due to its greater number of items. From a comparative standpoint then, the SES is not a much more sensitive measure than the SEI. If the SES were lengthened, not only might its reliability increase, but its overall sensitivity should be greater as well. Another way to increase the sensitivity of the SES would be to increase the number of possible responses for each item. This method however, may not prove very fruitful, as test takers may not be able to distinguish between “strongly agree”, “agree”, and “agree somewhat”, for example.

The biggest problem facing the SES may be its face validity. In many cases (e.g., ability testing), a test with high face validity is helpful. In these cases, having a test with high face validity can promote cooperation by the test taker. Tests with low face validity are less likely to be taken seriously, and may result in the test taker either refusing to take the test, or putting less effort into it. In addition, face validity is useful in legal situations. A face valid test is more likely to be accepted in courts and in law making decisions

(Anastasi & Urbina, 1997). In the case of the SES however, high face validity could be a detriment to the test's true validity. More specifically, the high face validity of the SES could influence responding of SES test takers. This phenomenon is a common problem among self-report personality tests. Test takers may intentionally or unintentionally give a distorted image of themselves when responding to these types of measures (Cohen & Swerdlik, 1999). This distortion relates to the previously mentioned concept of social desirability, which was noted as a possible confound for the Coopersmith SEI. To be sure, many of the problems found with the SEI are most likely to be found with the SES as well, since both tests measure self-esteem in a similar fashion.

Snuffer (2004) found the face validity of the SES to be quite high. Face validity was assessed by asking participants ($N = 30$) to examine a copy of the SES (without being told what it was), then to write down, on the back of the test, what the test appeared to be measuring. Of the 30 participants tested, 21 (70%) specifically stated that self-esteem was being measured. Of the nine remaining responses, five were judged by a clinician as being equivalent to self-esteem. Thus, it was understood by 26 out of the 30 participants (86.66%) that the SES measured self-esteem.

Rosenberg (1965) addressed the issue of face validity concerning his test. While he acknowledges that there is “little doubt” (p. 18) that the test items deal with self-esteem, he also argues that face validity is not an issue. He wrote, “...if this scale actually did measure self-esteem, then we would expect the scores on this scale to be associated with other data in a theoretically meaningful way” (p. 18). Rosenberg then backed his argument by showing test scores were associated with depression.

In his study, Rosenberg tested 50 “normal” volunteers housed in a clinical center.

The 50 participants were asked to complete the SES, as well as other questionnaires. Nursing staff at the center were asked to describe the participants, either yes or no, as being “often gloomy.” It should be noted that the participants were not aware that the nurses were judging them, nor were the nurses aware that the participants had taken a self-esteem test. Rosenberg found that 21 of the 23 (91%) high self esteem scorers and 12 of the 15 (80%) medium self-esteem scorers were described by the nurses as not “often gloomy.” Only 3 of the 12 (25%) low self-esteem scorers were described by the nurses as not “often gloomy.”

At a first glance, it would appear that there is a clear connection between SES scores and gloominess, or depression. However, some important flaws need to be considered. First, and perhaps most importantly, Rosenberg never mentions what scores would be required to be classified as having either high, medium, or low self-esteem. Secondly, as a related point, there were an uneven number of participants in each of the three classifications. One can only wonder if Rosenberg defined his classifications post hoc, putting the top 23 self-esteem scorers in the “high” group only because the top 16 scorers (i.e., the top third) did not make his point as clearly. Conversely, if the “low” scorers had consisted of 16 participants instead of only 12, and the four highest self-esteem scorers of this group were judged as not “often gloomy,” that would mean that 44% of low scorers, as opposed to 25% found in Rosenberg's study would be labeled as not “often gloomy.” By comparison, only 25% of low scorers would be labeled as “often gloomy,” and 31% would be labeled as “undecided.” The point is that the inequality of the group sizes would be understandable if classification criteria were listed. However, Rosenberg does not provide these criteria. How then, can one compare the groups, if the

groups are not clearly defined?

It also appears that for participants who were labeled as having high self-esteem, gloominess was certainly not a problem. For participants with low self-esteem however, the clarity of a problem with gloominess is not quite as apparent. While only 3 of the 12 (25%) low self-esteem scorers were labeled as not “often gloomy,” only 4 of the 12 (33%) low self-esteem scorers were labeled by the nurses as “often gloomy.” A re-analysis of Rosenberg's data shows that people with low self-esteem are as likely to be described as not “often gloomy” as they are to be described as “often gloomy,” $\chi^2(2, N = 12) = 4.34, p > .05$.

Another finding of Rosenberg (1965) was that participants with low self-esteem were more likely than participants with high self-esteem to present a façade to others. SES respondents were asked to agree or disagree with the statements, “I often find myself ‘putting on an act’ to impress people” and “I tend to put up a ‘front’ to people” (pp. 154-155). For both statements, lower self-esteem led to greater agreement. Once again, no cutoff criteria for high or low self-esteem are mentioned and a large discrepancy is found in the number of participants with low (38) versus high (484) self-esteem. Also, re-analysis of Rosenberg's data shows that low self-esteem scorers were as likely to disagree with both statements as they were to agree with one or both of the statements, $\chi^2(2, N = 38) = 2.96, p > .05$. Thus, not all persons with low self-esteem are likely to put on a façade. Another obvious problem with Rosenberg's façade study is the method he used to assess tendency to present a façade. Such a blatant method lends itself to low validity. It is quite possible, in fact probable, for a person who is likely to “put on an act”, whether they have high or low self-esteem, to disagree with either of the two

statements. People who agree with the two statements are either telling the truth, stating directly that they are likely to put on a façade, or lying, which would be a façade itself. People who disagree with the statements either truly are not likely to present a façade, or are lying and covering their tendency to present a façade with a façade. Essentially, it would be nearly impossible to tell exactly who is likely to present a façade simply based on the two questions that Rosenberg asked. Finally, if people with lower SES scores truly are more likely to present a façade, then how can one be sure that these people truly have low self-esteem? Low self-esteem scorers may simply be “faking bad.”

Despite these problems, to say that a test like the SES or the SEI is useless due to its high face validity is a misnomer. It is neither the point nor desire of the author to label these tests as such. If test takers have accurate insight into their own thinking, and are willing to respond honestly, self-report measures of self-esteem, or any other construct, can be very useful (Cohen & Swerdlik, 1999). The problem is that a test taker’s honesty may not be fully intact. It is believed that lowering the face validity of a self-report self-esteem measure can increase the likelihood of truthful responding, thereby making a more truly valid test for all.

Current Study

The current study sought to examine the effect of *intervening items* on scores of self-esteem, as measured by the SES. More specifically, what is the effect of placing a number of socially neutral items between items of the SES? While there appears to be no empirical information on this potential effect, it has been noted that neutral items may reduce socially desirable responding (Anastasi & Urbina, 1997). The underlying logic is that neutral items may “distract” a respondent from the overall theme of self-esteem

testing, creating a test in which the respondent cannot determine what is being measured. Several hypotheses were tested in this study. The first of these addressed the effect of intervening items on SES scores. The other hypotheses tested the validity of the SES.

Hypothesis 1. First, it was hypothesized that participants who were exposed to intervening items between items of the SES would have lower self-esteem scores than participants who responded to all 10 items of the SES together (i.e., without intervening items). It was believed that the intervening items would lower the face validity of the SES, which would decrease the likelihood of participants encountering these intervening items to “fake good”, resulting in lower SES scores.

Hypothesis 2. As stated earlier, self-esteem has been positively linked to erotic love and relationship satisfaction (Campbell, Foster & Finkel, 2002; Hendrick, Hendrick, & Adler, 1988). Therefore, it was hypothesized that participants in romantic relationships would have higher SES scores than participants who are not.

Hypothesis 3. Third, self-esteem scores should be positively correlated with the number of extracurricular activities in which a participant is involved. This hypothesis relates to Rosenberg's (1965) finding that adolescents with low self-esteem were less likely to participate in extracurricular activities, as well as Francis' (1997) finding that self-esteem is related to extraversion.

CHAPTER II

Method

Participants

The participants were 42 undergraduate students (19 males, 23 females) at Marshall University. The participants were sampled from introductory psychology courses at the university. Participation was voluntary, and participants were given extra course credit for their participation. The ages of the participants ranged from 18 – 32 years ($M = 19.14$ years). Participants who failed to respond to all items of the dependent measure were not included, but were given extra credit for their efforts.

Materials

Testing materials consisted of a 50-item questionnaire, a demographics survey, and a four item questionnaire assessing participants' self-perceived honesty and self-esteem. The 50-item questionnaire was comprised of the 10-item Rosenberg Self-esteem Scale (SES) and an additional 40 neutral items created by the experimenter. These 40 items were similar in format and nature to the SES, but differed from the SES in content. The 40 additional items, like the SES items, were statements referring to personal qualities to which the participants were asked to state to what degree they agree with the statements via a 4-point Likert scale ranging from strongly agree to strongly disagree. However, these 40 items were not directly related to self-esteem. The instructions for the 50-item questionnaire are shown in Appendix B.

The 50-item questionnaire was delivered in three forms, defining the three groups of the study. One form of the questionnaire involved participants first responding to the 10 items of the SES, followed by a question asking what those 10 items were about.

Afterwards, the 40 neutral items were presented. Participants who responded to this form of the questionnaire were classified as being in group ROSENBERG. A second form of the questionnaire involved participants responding to all 50 items of the questionnaire in a mixed order. That is, the 10 items of the SES were randomly included in the 40 neutral items. Block randomization was used, such that five items from the SES were used in each half of the questionnaire. After the 50 item questionnaire, a question asking what the items on the questionnaire were about was presented. Participants who responded to this form of the questionnaire were classified as being in group MIXED. To counterbalance for effects of fatigue, a third form was implemented. The third form of the questionnaire involved participants first responding to the 40 neutral items, followed by the 10 items of the SES, then a question asking what the items on the questionnaire were about. Participants who responded to this form of the questionnaire were classified as being in group REVERSE.

The demographics survey, which followed the 50-item questionnaire, asked participants for their age, gender, relationship status, religion, overall GPA, and to which clubs/groups the participant belongs. Also on this form, participants were asked whether they could identify any recurring themes among the items of the 50-item questionnaire, and if so, to list them. The demographics survey is shown in Appendix D

Following the demographics survey, participants were asked four questions related to their honesty and self-esteem. Participants were asked to rate, on a scale of 1 to 10, with 10 being most honest, how honest they were in their responding to the 50-item questionnaire. Next, participants were asked to rate how honest others would perceive them to be. That is, on a scale of 1 to 10, with 10 being most honest, how might another

person rate your honesty in everyday life? Participants were then asked to rate their own level of self esteem on a scale of 1 to 10 with 10 being highest, as well as what rating they felt others would give them concerning self-esteem. These four questions are listed in Appendix E.

Procedure

The experiment took place in a classroom setting. The experimenter explained the general nature of the experiment to the participants, stating that they would be taking a self-report personality test of which they would not know their individual results. Those who wished to participate were given two copies of an informed consent form to sign. After participants signed the informed consent form, the experimenter took one copy from each participant.

Participants were then given the 50-item questionnaire. The questionnaires were distributed using block randomization. Each block consisted of a ROSENBERG packet, a MIXED packet, and a REVERSE packet. After all participants completed their questionnaires, the questionnaires were individually collected by the experimenter. While collecting the questionnaires, the experimenter also distributed the demographics surveys and honesty/self-esteem questionnaires (DSHSQs). After all participants had completed the DSHSQ, the DSHSQs were collected by the experimenter. The experimenter then debriefed the participants and answered any questions that the participants had. The testing session lasted approximately 35 minutes.

CHAPTER III

Results

The mean SES scores for each experimental group were as follows. Mean SES score for group ROSENBERG was 30.50 ($SD = 3.59$). Mean SES scores for group MIXED was 31.23 ($SD = 4.27$). Mean SES scores for group REVERSE was 31.20 ($SD = 6.72$). For testing of the three hypotheses of this study, as well as any additional analyses, an alpha level of 0.05 was utilized.

Hypothesis 1

The first hypothesis was tested using a one-way, between-subjects ANOVA on SES scores which found no significant effect, $F(2, 39) = 0.091, p = 0.913$. Planned pairwise comparisons were conducted, with all pairwise comparisons being examined. No pairwise comparisons were significant using Bonferroni adjustments. A complex comparison examining the mean SES scores of both control groups (ROSENBERG and REVERSE) compared to the mean SES scores of group MIXED was conducted. No significant effect was found using Bonferroni adjustments, $F(1, 39) = 0.050, p > .05$.

Hypothesis 2

The second hypothesis was tested using an independent samples t -test. No significant difference was found between SES scores of participants who were not in romantic relationships ($n = 19, M = 30.84$) and those who were ($n = 23, M = 31.09$), $t(40) = -0.156, p = 0.877$.

Hypothesis 3

The third hypothesis was tested using a Pearson product-moment correlation. The correlation between SES scores and number of extracurricular activities was found to be

significant, $r(42) = 0.431, p < 0.01$.

Additional Analyses

Testing of the first hypothesis indicated that no significant difference exists between self-esteem scores of participants who respond to intervening items between items of the SES and participants who do not. To better understand why the obtained results were found, several follow-up tests were performed. These tests were conducted to examine possible confounding factors which may have led to either increased self-esteem scores in the mixed condition or lowered self-esteem scores in the control conditions.

To determine whether religion and subsequently, honesty related to religion may have been a factor in inflating SES scores, three independent samples t -tests were conducted. First, an independent samples t -test was conducted to determine whether religious participants had higher SES scores than non-religious participants. Mean SES score for participants claiming no religion ($n = 8$) was 27.63 ($SD = 5.88$). Mean SES score for participants claiming a religion ($n = 34$) was 31.76 ($SD = 4.51$). With equal variances not assumed, a marginally significant difference was found, $t(9) = -1.867, p = 0.095$.

Considering that religious participants had higher SES scores than non-religious participants, it is important to examine whether honesty related to religion could have been a factor. If it is found that religious participants are more honest, then lowered SES scores of non-religious participants could be indicative of “faking bad.” If it is found that religious participants are less honest, then elevated SES scores of religious participants could be indicative of “faking good.” If religious and non-religious participants are equal

in their levels of honesty, than elevated SES scores of religious participants would be due to factors other than honesty. To determine whether participants claiming a religion were more honest than participants claiming no religion, two independent *t*-tests were performed. The first compared religious and non-religious participants in self ratings of honesty (SELFHON). The second compared religious and non-religious participants in perceptions of their everyday honesty as would be rated by others (HONOTHER).

Mean SELFHON for religious participants ($n = 34$) was 9.53. Mean SELFHON for non-religious participants ($n = 6$) was 9.33. With equal variances not assumed, no significant difference was found, $t(6) = -.565, p = .593$.

Mean ratings of HONOTHER for religious participants was 8.91. Mean ratings of HONOTHER for non-religious participants was 8.33. With equal variances not assumed, no significant difference was found, $t(6) = -.917, p = .397$.

Regardless of their religion, if participants in the mixed condition are less honest than participants in the control groups, then it could be possible that elevated SES scores for group MIXED were due to “faking good,” which would indicate that the manipulation most likely had failed. To determine whether honesty may have influenced a particular group's score more than another group's score, two analyses of variance, as well as follow-up tests were conducted. First, a one-way, between-subjects ANOVA was conducted to examine the difference in SELFHON among the three experimental conditions. Mean SELFHON for group ROSENBERG was 9.69. Mean SELFHON for group MIXED was 9.17. Mean SELFHON for group REVERSE was 9.60. A marginally significant difference was found, $F(2, 37) = 3.019, p = 0.061$. Pairwise comparisons were conducted, with all pairwise comparisons being examined. No pairwise comparison

was found to be significant, using Tukey adjustments. Secondly, a one-way, between-subjects ANOVA was conducted to examine the difference in HONOTHER among the three experimental conditions. Mean HONOTHER for group ROSENBERG was 9.23. Mean HONOTHER for group MIXED was 8.33. Mean HONOTHER for group REVERSE was 8.87. A marginally significant difference was found, $F(2, 37) = 3.053$, $p = 0.059$. Pairwise comparisons were conducted, with all pairwise comparisons being examined using Tukey adjustments. Between groups ROSENBERG and MIXED, a significant difference was found. Between groups ROSENBERG and REVERSE, no significant difference was found. Between groups MIXED and REVERSE, no significant difference was found. Results of these pairwise comparisons are shown in Table 1.

Table 1

Pairwise Comparisons for Ratings of Honesty as Perceived by Others

<u>Groups</u>	<u>df</u>	<u>F</u>	<u>p</u>
ROSENBERG / MIXED	1, 37	6.057**	.048
ROSENBERG / REVERSE	1, 37	1.113	.548
MIXED / REVERSE	1, 37	2.286	.297

Note. ** $p < .05$

A paired samples t -test was conducted to examine the difference between participants' ratings of their own honesty on the questionnaire and their perceptions of their everyday honesty as would be rated by others. To the extent that self ratings of honesty were significantly higher than perceived everyday honesty, it could be inferred that participants found it easier to be honest while responding to the questionnaire than they do in everyday activities with others. This idea, which relates to task demands, will

be discussed later. Mean self honesty ratings were 9.50. Mean perceived honesty as rated by others was 8.83. A significant difference was found, $t(39) = 5.356, p < 0.01$. Paired samples t -tests were then conducted for each of the three experimental conditions to determine whether the difference between self-honesty ratings and honesty ratings as perceived by others varied as a function of condition. The difference did not vary by condition. Results of these tests are shown in Table 2.

Table 2

Paired Samples t -tests for self vs. others Honesty Ratings

Group	df	t	p
ROSENBERG	12	2.521**	.027
MIXED	11	3.079***	.010
REVERSE	14	3.556***	.003

Note. ** $p < .05$ *** $p \leq .01$

To establish the likelihood of a participant or group of participants to present a façade concerning self-esteem, difference scores were calculated, for each participant, between self ratings of self-esteem and ratings of perceptions of their self-esteem as would be rated by others. A paired samples t -test was then conducted to examine the difference between participants' ratings of their own self-esteem and their perceived self-esteem as rated by others. Mean ratings of self-esteem as perceived by self was 7.48. Mean ratings of self-esteem as perceived by others was 8.18. A significant difference was found, $t(39) = -3.445, p < 0.01$. Paired samples t -tests were then conducted for each of the three experimental conditions to determine whether the difference between self ratings of self-esteem and ratings of self-esteem as perceived by others varied as a

function of condition. For group ROSENBERG, no significant difference was found. For group MIXED, a significant difference was found. For group REVERSE, a marginally significant difference was found. Results of these tests are shown in Table 3.

Pearson correlations were conducted between SES scores, self-ratings of self-esteem (SELF-SE), and ratings of self-esteem as perceived by others (SE-OTHERS). All correlations were significant at $\alpha = 0.01$. Coefficients of these correlations are shown in Table 4.

Table 3

Paired Samples t-tests for self vs. others Self-Esteem Ratings

Group	<i>df</i>	<i>t</i>	<i>p</i>
ROSENBERG	12	-1.674	.120
MIXED	11	-2.462**	.032
REVERSE	14	-1.911*	.077

Note. * $p < .10$ ** $p < .05$

Table 4

Correlations Between SES Scores, SELF-SE, and SE-OTHERS

Variable	1	2	3
SES	--	.631***	.431***
SELF-SE		--	.654***
SE-OTHERS			--

Note. *** $p < .01$

Across groups, participants' self ratings of self-esteem were lower than their self-esteem as perceived by others, implying that participants may be presenting façades

regarding their self-esteem when around others. Presumably, the lower the difference score, the more likely a participant would be to present a façade. A one-way, between-subjects ANOVA, as well as follow-up tests were conducted to determine whether high scorers on the SES ($n = 13$), middle scorers on the SES ($n = 13$), and low scorers on the SES ($n = 14$) differed in this likelihood. These tests were conducted as a form of replication of Rosenberg's (1965) façade study, in which it was found that participants with low self-esteem were more likely to present a façade. Results similar to those found by Rosenberg would give support for the validity of the SES, assuming that low SES scorers truly are more likely to present a façade. High SES scorers (i.e., the top third of the present sample) were participants who scored higher than 34 out of a possible 40 on the SES. Middle SES scorers were participants whose scores ranged from 29 to 33 out of a possible 40. Low SES scorers were participants who scored lower than 29. Mean difference scores for low SES scorers were -0.714. Mean difference scores for middle SES scorers were -1.231. Mean difference scores for high SES scorers were -0.154. A marginally significant difference was found, $F(2, 37) = 2.45, p = .10$. Pairwise comparisons were conducted, with all pairwise comparisons being examined using Tukey adjustments. Between low and middle SES scorers, no significant difference was found. Between low and high SES scorers, no significant difference was found. Between middle and high SES scorers, a marginally significant difference was found. Results of these pairwise comparisons are shown in Table 5.

Additionally, a Pearson correlation was conducted between SES scores and self-esteem difference scores (i.e., tendency to present a façade). The correlation was not significant, $r(40) = 0.220, p = 0.172$.

Table 5

Pairwise Comparisons for Self-Esteem Difference Scores

SES Scorers	<i>df</i>	<i>F</i>	<i>p</i>
Low/Middle	1, 37	1.171	.531
Low/High	1, 37	1.378	.476
Middle/High	1,37	4.906*	.082

Note. * $p < .10$

To determine whether the 40 intervening items were truly neutral (i.e., did not measure self-esteem), Pearson correlations were conducted between each of the 40 intervening items and SES scores. Eight of the 40 items were significantly correlated with SES scores. A listing of the 40 intervening items, as well as their correlations with SES scores, can be found in Appendix C. Inter-item reliability for the SES was computed, and was high, $\alpha = 0.89$. This coefficient is similar to those found in previous studies (e.g., Gudjonsson & Sigurdsson, 2003). Inter-item reliability for the 10 SES items and the eight intervening items which correlated with SES scores was computed as well. The inter-item reliability for these 18 items was moderately high, $\alpha = 0.74$.

A qualitative analysis of responses given to the question, “Did you notice any recurring themes in the 50 items you responded to? If so, what were they?” was performed. Responses were arbitrarily judged by the experimenter whether the participant had identified self-esteem as a recurring theme. Frequencies are shown in Table 6. To determine whether a particular group was more likely to identify self-esteem as a recurring theme in the 50-item questionnaire, a one-tailed chi-square was conducted. No significant effect of group was found on participants' ability to name self-esteem as a

recurring theme of the questionnaire, $\chi^2(2, N = 42) = 2.98, p > .05$. However, in a comparison of only groups ROSENBERG and MIXED, a marginally significant difference was found, $\chi^2(1, N = 27) = 2.97, p = 0.085$.

Table 6

Frequencies of Identifying Self-Esteem as a Recurring Theme by Group

<u>Group</u>	<u>Identified Self-Esteem</u>		
	No	Yes	Total
ROSENBERG	4	10	14
MIXED	8	5	13
<i>REVERSE</i>	7	8	15

CHAPTER IV

Discussion

Analysis of the data suggests that intervening items do not affect SES responding, and that the SES is a valid test. Supporting this notion is the finding that SES scores were positively and significantly correlated with extracurricular activities. However, contrary to earlier findings (e.g., Campbell, et al, 2002), SES scores were not affected by the presence of a romantic relationship. While participants in relationships did have higher SES scores than participants not in relationships, the difference between these groups was not significant. This finding may be due to a combination of two factors. First, in earlier research, the correlations between SES scores and relationship status, while significant, were rather low. Secondly, considering that the aforementioned correlations were low, it is possible that the sample for the current study was too small to detect significance. The earlier studies conducted by Campbell, et al (2002) and Hendrick, et al (1988) utilized samples of 138 participants and 114 participants, respectively. The sample used in the current study consisted of only 42 participants. It should also be noted however, that the previous studies mentioned examined the correlations between self-esteem and relationship satisfaction, not merely the presence of a relationship, which was examined in the current study. Therefore, it is possible that for people in relationships, self-esteem is related to their perception of the relationship, though the presence of their relationship does not influence self-esteem.

While it is possible that neutral intervening items truly have no effect on SES responding, alternate explanations for the results are available. Some of the possible explanations for the results obtained are task demands and non-neutrality of the

intervening items. These possible confounds, as well as other issues, should be discussed.

The rationale behind the experiment assumes that the participants may not have total honesty in their responding. However, if the sample, as a whole, is honest in their responding, it becomes more likely that the self-esteem scores of each group should not differ significantly as a function of the manipulation. The purpose of the manipulation was to make some test takers (i.e., group MIXED) more honest in their responding, thereby lowering their tendency to “fake good,” resulting in lowered SES scores for that group. However, if the sample is already honest, the tendencies of participants in group ROSENBERG to “fake good” should already be lowered, and any effect of the intervening items would be obscured. An examination of the honesty of the sample is therefore important.

Over the course of the present study, it became theorized that honesty, as a function of religion, may have played a role in the observed results. That is, religious participants, who presumably are more honest, should be less likely to “fake good” and thereby have lower SES scores. This hypothesis developed due to the fact that a large proportion of the sample, both within and across groups, claimed to have a religion. In addition, the proportion of religious participants in groups ROSENBERG and REVERSE (85.7% and 80%) respectively is higher than in group MIXED (76.9%). The higher proportion of religious participants in groups ROSENBERG and REVERSE could suggest that those participants were consequently less likely to “fake good,” which could help explain why their SES scores were no higher than participants in group MIXED. However, it was found that religious participants are not significantly more honest than non-religious participants. Any differences between religious and non-religious

participants in terms of self-esteem, then, should not be attributed to higher honesty on the part of the religious participants. Therefore, while it was shown that the presence of religion may elevate a person's self-esteem, it is doubtful that elevation in SES scores in the mixed group (as compared to control groups) can be explained by honesty (or dishonesty) related to religious factors.

A more likely factor influencing the honesty of the sample may have been anonymity. Participants were aware, at the time of the study, that the experimenter would not be able to recognize any individual's scores on the questionnaire. In addition, participants were tested as a group, which may have added to feelings of security among the participants by limiting the personal interaction between participants and the experimenter. Therefore, the participants had little or no reason to withhold the truth. This notion is supported by the fact that participants, across and within experimental conditions, rated their responding to the 50-item questionnaire as being more honest than their perceptions of their everyday honesty around others. Additionally, no difference was found between the three groups in terms of self-reported honesty on the 50-item questionnaire. This finding implies that all participants, regardless of condition, may have found it easier to be honest for a short period of time in an anonymous situation than they do in everyday life. Task demands, then, may have influenced the results. Different results may have occurred if the participants had been tested on an individual basis, especially in a clinical setting, where the SES is more commonly utilized. Research should be conducted using individual testing, to determine if these effects play a role. The implications of a study of this nature would hold greater relevance for clinical uses of the SES.

Another possible influence on the honesty of the sample may have been instructions by the experimenter. All participants were instructed (as shown in Appendix D) to “answer truthfully.” It is possible that these instructions implied to the participants that honesty was more essential than usual, which could explain why self ratings of honesty concerning the 50-item questionnaire were higher than ratings of everyday honesty as perceived by others.

One of the assumptions of the experimental manipulation was that the 40 intervening items would be truly neutral, or unrelated to self-esteem. Any items that violated this assumption could undermine the effectiveness of the experimental manipulation.

As mentioned before, 8 of the 40 intervening items were significantly correlated with SES scores. Coefficient alpha for these eight items in combination with the SES was moderately high ($\alpha = 0.74$), indicating that all 18 items were measuring self-esteem. It is possible that participants were influenced by these eight intervening items, maintaining thoughts related to self-esteem, and adjusting their responses accordingly, despite reporting that they were being honest. Concordantly, chi-square analysis showed that there was no significant effect of group on participants’ ability to name self-esteem as a recurring theme of the questionnaire. Also important to note is that for all participants, a majority of the participants (54.8%) were able to identify self-esteem as a recurring theme of the 50-item questionnaire. These findings indicate that participants in group MIXED were as likely as participants in the two control groups to identify a theme of self-esteem in the questionnaire and suggest that the manipulation was not as strong as expected. Any replication of the current study should involve replacement of the eight

items that correlate with the SES with eight new items that do not correlate with Rosenberg's scale.

Another related possibility could be that participants assumed from the beginning that some items on the questionnaire would be related to self-esteem. Participants were told before receiving the questionnaire that they would be taking a “personality test.” As a result, some participants may have anticipated self-esteem items by making a connection between personality and self-esteem. That is, the participants assumed that because self-esteem is part of personality, a personality test should have items related to self-esteem.

To this point, honesty has been discussed as a factor influencing all three experimental conditions. While this discussion could partly explain why no significant difference was found between the three groups in terms of SES scores, it does not begin to explain why participants in the mixed condition had higher SES scores than participants in the two control conditions. Some of the findings of the current study suggest that participants in the mixed condition were more likely to present a façade, which may explain elevated SES scores. Likelihood to present a façade was assessed by calculating difference scores between self ratings of self-esteem and ratings of self-esteem as perceived by others. Negative difference scores indicated likelihood to present a façade. This likelihood would increase as difference scores became more negative. In all three of the experimental conditions, mean difference scores were negative. However, results of paired samples *t*-tests showed that only for group MIXED was this difference significant. Therefore, while some confounding factors may have led to overall high levels of honesty among the participants, other factors may have decreased the honesty in

group MIXED. It is possible, then, that the experimental manipulation not only failed but increased a participant's tendency to give socially desirable responses. The aforementioned increase may have been caused by greater attention being given by group MIXED to the content of the intervening items, as the theme of self-esteem was present, yet inconsistent in the mixed condition.

Consistent with the greater likelihood of participants in group MIXED to present a façade are results of the one-way ANOVA examining differences in honesty among the three groups. In ratings of honesty as perceived by others (though not in self ratings of honesty in responding to the 50-item questionnaire), participants in group MIXED rated themselves as less honest than the group ROSENBERG.

What appears to be the most important finding of the current study is that people, whether they score high or low on the SES, seem to have tendencies to present façades when it comes to their self-esteem. This assertion is evidenced by the findings that a significant difference exists between participants' self ratings of self-esteem and their ratings of self-esteem as perceived by others, and that the correlation between SES scores and difference scores (i.e., tendency to present a façade) was insignificant. Also, while SES scores were significantly correlated with both self-ratings of self-esteem and ratings of self-esteem as perceived by others, the correlation was stronger for self-ratings of self-esteem. This stronger correlation suggests that a better predictor of self-esteem is how people perceive their own self-esteem, as opposed to the perceived level of self-esteem people portray to others, if the SES truly measures self-esteem.

Assuming that the aforementioned assertion is true (i.e., people tend to present façades concerning self-esteem), this tendency may cause problems in assessing self-

esteem whether in anonymous situations, such as this study, or in individual/clinical situations, where self-esteem, as well as countless other personality states and traits, become a more sensitive issue. The personal interaction between a patient (i.e., a test taker) and a clinician may influence patient responding such that a patient would be more likely to present a façade, to hide the fact that something is wrong. Given the stigma that often follows mental health patients, behavior that would lower chances of a diagnosis would not be surprising.

Ultimately then, the question underlying the present study still remains: Does the SES measure self-esteem, or does it fall prey to socially desirable responding? The answer comes in the form of more questions. Does the SES accurately measure self-esteem in all people, or more so for honest ones? Is an anonymous situation required to ensure honesty? These questions, and indeed many others, stem from the findings of the current study. Further research is required to answer these questions, as well as to ensure that measurement of this important construct is being done appropriately.

References

- Anastasi, A. & Urbina, S. (1997). *Psychological Testing* (7th ed.). New Jersey: Prentice- Hall.
- Campbell, W.K., Foster, C.A., & Finkel, E.J. (2002). Does self-love lead to love for others? A story of narcissistic game playing. *Journal of Personality and Social Psychology*, 83, 340-354.
- Cohen, R.J., & Swerdlik, M.E. (1999). *Psychological Testing and Assessment: An Introduction to Tests and Measurement* (4th ed.). Mountain View, CA: Mayfield.
- Coopersmith, S. (1967). *The Antecedents of Self-Esteem*. San Francisco, CA: Freeman.
- Cullen, D. (1997). Maslow, monkeys, and motivation theory. *Organization*, 4, 355-373.
Retrieved June 11, 2004, from
<http://www.ualberta.ca/~holson/589/cullen/maslow.htm>
- Dechesne, M., Pyszczynsky, T., Arndt, J., Ransom, S., Sheldon, K.M., van Knippenberg, A., et al. (2003). Literal and symbolic immortality: The effect of evidence of literal immortality on self-esteem striving in response to mortality salience. *Journal of Personality and Social Psychology*, 84, 722-737.
- de Man, A.F., & Gutierrez, B.I.B. (2002). The relationship between level of self-esteem and suicidal ideation with stability of self-esteem as moderator. *Canadian Journal of Behavioural Science*, 34(4), 235-238.
- Fleming, J.S., & Courtney, B.E. (1984). The dimensionality of self-esteem: II. Hierarchical facet model for revised measurement scales. *Journal of Personality and Social Psychology*, 46, 404-421.
- Francis, L.J. (1997). Coopersmith's model of self-esteem: Bias toward the stable

- extravert? *The Journal of Social Psychology*, 137, 139-142.
- Goldney, R.D. (2003). Depression and suicidal behavior: The real estate analogy. *Crisis*, 24, 87-88.
- Gudjonsson, G.H., & Sigurdsson, J.F. (2003). The relationship of compliance with coping strategies and self-esteem. *European Journal of Psychological Assessment*, 19, 117-123.
- Harmon-Jones, E., Simon, L., Greenberg, J., Pyszczynski, T., Solomon, S., & McGregor, H. (1997). Terror management theory and self-esteem: Evidence that increased self-esteem reduces mortality salience effects. *Journal of Personality and Social Psychology*, 72, 24-36.
- Hendrick, S.S., Hendrick, C., Adler, N.L. (1988). Romantic relationships: Love, satisfaction, and staying together. *Journal of Personality and Social Psychology*, 54, 980-988.
- Hoyle, R.H., Kernis, M.H., Leary, M.R. & Baldwin, M.W. (1999). *Selfhood: Identity, esteem, regulation*. Boulder, CO: Westview Press.
- James, W. (1890). *The principles of psychology* (vol. 1). New York: Holt.
- Leary, M.R., Cottrell, C.A., & Phillips, M. (2001). Deconfounding the Effects of dominance and social acceptance on self-esteem. *Journal of Personality and Social Psychology*, 81, 898-909.
- Leary, M.R., Tambor, E.S., Terdal, S.K., & Downs, D.L. (1995). Self-esteem as an interpersonal monitor: The sociometer hypothesis. *Journal of Personality and Social Psychology*, 68, 518-530.
- Manis, M. (1955). Social interaction and the self concept. *Journal of Abnormal and*

- Social Psychology*, 51, 362-370.
- Maslow, A.H. (1940). A test for dominance-feeling (self-esteem) in college women. *Journal of Social Psychology*, 12, 255-270.
- Measure: Rosenberg self-esteem scale (SES; n.d.). Retrieved June 16, 2004, from <http://eib.emcdda.eu.int/Studies/arf-ses.htm>
- Morretti, M.M. & Higgins, E.T. (1990). Relating self-discrepancy to self esteem: The contribution of discrepancies beyond actual-self ratings. *Journal of Experimental Social Psychology*, 26, 108-123.
- Pearl, D. (1954). Ethnocentrism and the self concept. *Journal of Social Psychology*, 40, 138-147.
- Pelham, B.W. (1995). Self-investment and self-esteem: Evidence for a Jamesian model of self-worth. *Journal of Personality and Social Psychology*, 69, 1141-1150.
- Peterson, C., & Austin, J.T. (1985). [Review of the test Coopersmith Self-Esteem Inventories]. In J.V. Mitchell, Jr. (Ed.), *The ninth mental measurements yearbook* (pp. 395-398). Lincoln, NE: Buros Institute of Mental Measurements.
- Reasoner, R. (2004). *The true meaning of self-esteem*. Retrieved June 10, 2004, from National Association for Self-Esteem Web site: <http://www.self-esteem-nase.org/whatisselfesteem.shtml>
- Reckless, W., Dinitz, S., & Kay, B. (1957). The self component in potential delinquency and potential non-delinquency. *American Sociological Review*, 22, 566-570.
- Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.
- Sewell, T.E. (1985). [Review of the test Coopersmith Self-Esteem Inventories]. In J.V.

- Mitchell, Jr. (Ed.), *The ninth mental measurements yearbook* (pp. 395-398).
Lincoln, NE: Buros Institute of Mental Measurements.
- Silverstone, P.H., & Salsali, M. (2003, February 11). Low self-esteem and psychiatric patients: Part I – The relationship between low self-esteem and psychiatric diagnosis. *Annals of General Hospital Psychiatry*, 2, Article 2. Retrieved June 15, 2004, from <http://www.pubmedcentral.nih.gov/picrender.fcgi?arted=151271&action=stream&blobtype=pdf>
- Smith, H.M., & Betz, N.E. (2002). An examination of efficacy and esteem pathways to depression in young adulthood. *Journal of Counseling Psychology*, 49, 438-448.
- Snuffer, D. (2004). [Assessment of face validity of the Rosenberg SES]. Unpublished raw data.
- Ward, S. (1996). Filling the world with self-esteem: A social history of truth making [Electronic version]. *Canadian Journal of Sociology*, 21, 1-23.
- Watson, D., Suls, J., & Haig, J. (2002). Global self-esteem in relation to structural models of personality and affectivity. *Journal of Personality and Social Psychology*, 83, 185-197.

Appendix A

Items Of The Rosenberg Self-Esteem Scale (SES)

On the whole, I am satisfied with myself.

At times I think I am no good at all.

I feel that I have a number of good qualities.

I am able to do things as well as most other people.

I feel I do not have much to be proud of.

I certainly feel useless at times.

I feel that I'm a person of worth, at least on an equal plane with others.

I wish I could have more respect for myself.

All in all, I am inclined to feel that I am a failure.

I take a positive attitude toward myself.

Appendix B

Instructions For 50-Item Questionnaire

Below is a list of statements dealing with your general feelings about yourself. If you strongly agree, circle SA. If you agree with the statement, circle A. If you disagree, circle D. If you strongly disagree, circle SD. There are no correct or incorrect answers to these items. Just answer truthfully about yourself.

Appendix C

List Of Intervening Items And Their Correlations With The SES

I like the color red.	.091
I enjoy the sensation of being dizzy.	.068
I enjoy reading in my spare time.	-.007
I still have regular contact with my parents.	.202
I “bend over backwards” for the people that I care about.	-.063
The idea of a “higher power” seems foolish to me.	-.251
I sometimes smoke to relieve stress.	-.206
I like working with money.	.264
I often feel overwhelmed by deadlines.	-.410**
The idea of losing a family member terrifies me.	.050
When a friend shows me a new acquisition, I usually act impressed.	.177
I usually remember my dreams.	.240
I sometimes initiate conversations with strangers.	.154
I often worry about what other people think about me.	-.281
I enjoy outdoor activities.	.057
My career goals have changed over the last five years.	-.040
Being instructed to do something constantly adds stress to my life.	-.408**
I like to wear bright-colored clothing.	.176
I like to watch cartoons.	-.132
The idea of being famous is appealing to me.	.393**
I do not enjoy mathematics.	-.072

Being in control scares me.	- .564**
I am obsessed with being on time.	.580**
I am usually stressed out about finances.	-.348**
I find it easier to sleep with music turned on.	-.270
I often feel that I am being watched.	-.197
I feel comfortable in heavy traffic.	.316**
If someone in authority asks me to do something, I always obey.	.052
I usually have to search for my keys.	-.094
I am quite comfortable with being at home alone.	.278
I like the color blue.	.130
I would rather be a leader than a follower.	.476**
I feel that good looks are important for success.	-.062
If a friend had more of something than I did, I would be jealous.	-.204
I usually change clothes more than once a day.	.110
I prefer to be physically active, whenever possible.	.084
I usually take a nap sometime during the day.	.110
My best friend is of the opposite sex.	.212
I enjoy celebrating holidays.	.117
I give my help to others even if it puts myself out.	.185

Note. ** Correlation is significant at .05 alpha level

Appendix D

Demographics Survey

Did you notice any recurring themes in the 50 items you responded to? If so, what were they?

Please fill out the following information

Age _____

Gender: M F

Relationship Status:

____ single/not dating ____ dating ____ engaged ____ co-habiting

____ married ____ separated ____ divorced ____ widowed

What is your religion?

____ Agnostic/Atheist

____ Buddhist

____ Catholic

____ Hindu

____ Judaism

____ Muslim

____ Protestant

____ Other _____

What is your overall GPA? Use High School if necessary. _____

What extracurricular activities do you have? List them below.

Appendix E

Questions assessing self ratings of honesty and self-esteem

- 1) On a scale of 1 to 10, with 10 being most honest, how honest were you in your responses to items 1 – 50?
- 2) If a loved one or close friend were to rate your overall everyday honesty on a scale of 1 to 10, with 10 being most honest, how would they rate you?
- 3) On a scale of 1 to 10, with 10 being the highest, what do you feel is your level of self-esteem?
- 4) If a loved one or close friend were to rate your self-esteem on a scale of 1 to 10, with 10 being the highest, how would they rate you?