This is to certify that the
dissertation entitled

THE DISTORTING EFFECT OF EXCESSIVE DEFENSIVENESS ON SELF-REPORT MEASURES OF MENTAL HEALTH AND ON TREATMENT OUTCOME MEASUREMENT

presented by

BENJAMIN C. ADDLESON

has been accepted towards fulfillment
of the requirements for the

Doctoral degree in Psychology

Major Professor's Signature

4/9/04

Date

THE DISTORTING EFFECT OF EXCESSIVE DEFENSIVENESS ON SELF-REPORT MEASURES OF MENTAL HEALTH AND ON TREATMENT OUTCOME MEASUREMENT

By

Benjamin C. Addleson

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Psychology

2004

ABSTRACT

THE DISTORTING EFFECT OF EXCESSIVE DEFENSIVENESS ON SELF-REPORT MEASURES OF MENTAL HEALTH AND ON TREATMENT OUTCOME MEASUREMENT

By

Benjamin C. Addleson

The purpose of the present study was to examine whether the validity of self-report measures of mental would be brought into question by the effects of psychological defenses on responses to such measures in a treatment study. Although defensiveness did not moderate the relationship between self-report mental health and observer-rated mental health in this study, defensiveness was found to attenuate differences in self-report between time one and time two for the treatment group. Thus, there was some support for the hypothesis that taking defensiveness into account may yield more robust results in the treatment-outcome literature. The possible impact of non-pathological regression and ego resilience on results is discussed, as well as the ways in which the observer rated mental health measure – the Social Cognitions and Object-Relations Scale (Westen, 1990) may or may not have reflected such factors in the non-psychiatric, college population used in this study.

TABLE OF CONTENTS

TABLE OF CONTENTS
(Continued)

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Overview of the Argument

This study examined whether the validity of self-report measures of mental health[1] is reduced by the effects of excessive psychological defensiveness on them. A normal degree of defensiveness is thought to protect a person from consciously experiencing excessive anxiety or guilt (Westen, 1998). However, excessive defensiveness can be maladaptive if it distorts reality to a large degree (Cramer, 1991), because such distorted information processing can leave one vulnerable to any range of psychopathologies (Allen, Hauser, & Borman, 1996). Arguably, when people deny psychopathology to an unusual degree on self-report mental health questionnaires, we do not know to what degree such reports are accurate without accounting for the effect of psychological defenses.

This denial of psychopathology is a particularly thorny issue for treatment outcome measurement[2]. An effective treatment should decrease patients' excessive defensiveness and this should improve their ability to adapt to reality. As an individual becomes less defensive they function more adaptively, but are also likely to become more aware of painful or anxiety provoking thoughts and feelings to a degree. Unfortunately, this could translate into an individual endorsing more psychopathology *after* an effective intervention than *before* the intervention, even though they are functioning better.[3] Thus,

---

[1] In this paper I use the terms "self-report measures of mental health" and "self-report measures of psychopathology" synonymously as any difference between the two is not essential to the argument.

[2] The term "denial of psychopathology" in this paper refers simply to a failure to endorse psychopathology that is present. It does not refer specifically to the psychoanalytic meaning of denial, which will be referred to here as "disavowal," except where sited authors have used the term psychoanalytically.

[3] The argument that a lack of awareness of painful thoughts or feelings leads to better psychological health ("ignorance is bliss"), is not supported empirically when a lack of awareness is excessive, as will be discussed later. Briefly, increased awareness of thoughts and feelings may yield increased pain, but it also

using self-report of symptoms as the only measure of mental health or treatment outcome may leave us in the regrettable position of relying on a type of report that may not be valid because it does not account for the effects of psychological defenses.

Although some test authors have argued that their self-report measures should not be used in isolation (Eysenck, 1994), thousands of studies are relying on self-report measures as the sole measures of mental health (Shedler, Mayman, & Manis, 1994). A literature search conducted by the present author revealed that over the last 10 years, more than 100 published studies of psychotherapeutic efficacy or effectiveness, in particular, have used either the State-Trait Anxiety Inventory (Speilberger, 1983) or the Beck Depression Inventory (Beck, 1996) – self-report measures -- as the *only* gauge of mental health before and after an intervention.[4] None of the above efficacy or effectiveness studies empirically addressed the possibility of inaccurate self-report due to defenses such as disavowal, repression, or suppression.[5] This lack is despite substantial evidence suggesting that excessive defensiveness exists, has a negative effect on mental and physical health, and might have a distorting effect on self-report of mental health, as will be reviewed later.

Much of the current empirical literature on defensiveness can be divided into two related lines of research. One line of research has attempted to classify defenses according to their developmental maturity and adaptiveness. Another has focused on the

---

yields more data on which to make better choices. High levels of defensiveness are also related to other indices of poor mental and physical health. Of course, individuals may exist who are *not defensive enough.* In this case, effective treatment would increases defensiveness, though this process is not the primary issue in this paper, as a paucity of defenses should not cause an individual to deny psychopathology on self-report measures of mental health.

[4] This Literature search was onducted with PSYCH INFO using "Sate-Trait Anxiety Inventory and effectiveness or efficacy" and "Beck Depression Inventory and effectiveness or efficacy" as search terms. This was not intended to be an exhaustive search, but even such a preliminary search demonstrates the widespread usage of these self-report measures without accounting appropriately for defensiveness.

degree of defensiveness, that would generally fall under the rubric of immaturity. The former line of research will be termed research on "maladaptive defenses" for the purposes of this paper. The latter line of research will be termed "excessive defensiveness" research. This will be an important distinction to make later.

Although research on excessive defensiveness represents most of the empirical work on physiological and psychological costs of defensiveness, it lacks specificity in that it defines defenses (e.g, defensive style) poorly or not at all. Research on maladaptive defenses defines specific defenses, but suffers from reliance on self-report or on observer-report measures that either lack reliability or are too unwieldy to be practical in clinical use and replication efforts.[6] Whereas both lines of research have theoretical and empirical problems, each has contributed to a body of data suggesting that psychological defenses exist, can have serious mental and physical health costs (when excessive or immature), and may distort self-report of psychopathology.

Physiological data provide one kind of evidence supporting the theory of defensive operations. Individuals who are judged to be highly defensive by clinicians and measures of excessive defensiveness evidence physiological indices of stress such as higher heart rate and blood pressure (Shedler, Mayman, & Manis, 1993; Weinberger, Schwartz, & Davidson, 1979) as well as greater galvanic skin conductance (Dozier & Kobak, 1992; Hughes, Uhlman, & Pennebaker, 1994; Pennebaker, 1997; Weinberger et

---

[5] A list of defenses mentioned in this paper and their definitions is available in appendix A.

[6] Excessive defensiveness research also, technically, relies on self-report defensiveness measures, but they are quasi-projective in nature. That is, defensive individuals are expected to *deny* socially undesirable behaviors that are nearly universal on such measures (items such as "there have been occasions when I have felt like smashing things"). Defensive indiciduals are not expected to *endorse* items such as "I suppress feelings that are uncomfortable to me" on excessive defensiveness measures. Items such as the latter are, however, included on self-report measures of maladaptive defenses. Because of this difference between excessive defensiveness measures and maladaptive defenses measures, excessive defensiveness measures will not be termed "self-report" measures in this study.

al., 1979) in comparison to normatively-defensive controls (and frequently in comparison to those who endorse excessive anxiety) when undergoing a task designed to tap their anxieties (and, therefore, defenses). Another line of research has demonstrated that individuals thought by clinicians to have a reason to defend against the awareness of sexual arousal deny experiencing arousal in self-report, but demonstrate arousal physiologically (via plethysmograph) when exposed to arousing stimuli (Adams, Wright, & Lohr, 1996; Morokoff, 1985).

There is also empirical evidence suggesting that defensiveness is associated with poor psychological and physical health. Immature defenses correlate with poor global functioning, psychological illness and the presence of personality pathology (Perry & Cooper, 1989; Vaillant & Drake, 1985). Pennebaker and colleagues have found that the failure to express painful feelings or discuss traumatic events correlated with current and long-term health problems (Pennebaker, 1985; Pennebaker & Susman, 1988). They also found that writing about or discussing painful or traumatic events was associated with fewer health problems and health center visits, increased immune functioning, and more adaptive behavior. (Pennebaker, 1989; Pennebaker & Beall, 1986; Pennebaker & Heeron, 1984; Pennebaker, Glaser, & Glaser, 1988).

Independent lines of research have found similar results as those found by Pennebaker's group. Shedler et al., (1993) found that defensive denial was linked to large spikes in heart rate and blood pressure. These same increases in heart rate and blood pressure have been associated with damage to the inner lining of coronary arteries, leading to arteriosclerosis and heart disease (Krants & Manuck, 1984; Manuck, Kaplan & Clarkson, 1983). Jensen (1987) found that repressive personality style and reduced

expression of negative affect predicted the rate of breast cancer spread when the type and severity of cancer at the outset was controlled for. In short, there is a reasonable body of research suggesting that defenses exist and that there can be substantial mental and physical health costs for reliance on maladaptive or excessive defenses.

Turning back to the topic at hand, there is some evidence to suggest that such defenses distort people's report of their own psychopathology as well. Repressors (those who avoid the recognition or experience of negative affect) report low anxiety on the Taylor Manifest Anxiety Scale (Taylor, 1953) despite physiological and behavioral evidence suggesting they are, in fact, anxious (Davis, 1987; Mikulincer & Orbach, 1995; Weinberger et al., 1979). One study has found that defensive deniers report a lack of symptoms on the Eysenck Neuroticism Scale (Eysenck & Eysenck, 1975) and the Beck Depression Inventory (BDI, Beck, Ward, Mendelson, Mock, & Erbaugh, 1961), despite both physiological evidence and clinical judgment suggesting otherwise (Shedler, Mayman, & Manis, 1993). In fact, Beck, himself (Kendall, Hollon, Beck, Hammen, & Ingram, 1987) as well have others (Scafidi, Field, Prodromidis, & Abrams, 1999) have indicated that scores of 0 on the BDI are more likely indicators of a tendency to deny psychopathology than of a total lack of depressive symptomotology.

Research suggesting that defenses distort self-report measures of psychopathology and are associated with mental and physical health problems brings up the question of whether self-report measures of mental health are adequate as the sole measure of effectiveness or efficacy of a psychotherapeutic intervention. If defensiveness has a tendency to increase self-report ratings of mental health, then the results of effectiveness and efficacy studies could be attenuated by individuals who defensively denied pathology

at pretest and do so to a lesser degree at post-test. That is, although these participants may have experienced mental health gains, self-report measures of mental health might reflect little if any change, or even "deterioration" over the course of an intervention. Thus, we might be able to enhance our ability to interpret self-report mental health measures by taking defenses into account.

One goal of this study was to determine if excessive defensiveness would cause people to deny pathology on self-report questionnaires that was present by observer codes. However, a non-self-report measure of mental health was necessary for the present study to verify that an individual who denied pathology and appeared excessively defensive on a measure thereof was, by other measures, not as mentally healthy as they had reported. One might have found observer-rated scales such as the Hamilton Rating scale (Hamilton, 1960) or the Global Assessment Scale (Endicott, Spitzer, Fleiss, & Cohen, 1976) to be useful criterion measures of mental health. However, such a rating may still have been susceptible to distorting effects of defensiveness to the extent that defensiveness could hide outward signs of pathology to all but the most experienced clinicians utilizing an extended interview (Shedler et al, 1994). Also, conducting such extended interviews would have been exceedingly costly. Thus, in addition to observer ratings, a projective approach may have made a criterion measure of mental health less susceptible to distorting effects of defensiveness. This could be the case because the portrayal of one's psychopathology through a projective test purportedly occurs outside of awareness and therefore may not activate defensive functioning to the same degree as direct questions regarding psychopathology.

The Social Cognitions and Object Relations Scale (SCORS, Westen, 1991) is one example of such a measure that relies on assessment of object relations, using observer coding of narrative data, within a format that is not extremely time consuming to administer. It has been judged to be more valid and reliable than several other observer-rated projective measures (Lilienfeld et al., 2000).

It was hypothesized that self-report mental health questionnaires (the BDI and STAI being examples) would be increasingly inaccurate at higher levels of self-reported health due to the distorting effects of maladaptive defenses. It was further hypothesized that self-report mental health measures would not accurately reflect improvement after a psychotherapeutic intervention. It was expected that excessive defensiveness would account for this inaccuracy insofar as decreased defensive functioning would account for attenuated scores on self-report mental health measures.

The following literature review describes early historical perspectives on defenses that have shaped current work on excessive defensiveness and maladaptive defenses. This is followed by a review of literature that has provided evidence suggesting the existence of excessive defensiveness, its effects on psychological and physical health, and the possibility that it has a distorting effect on self-report questionnaires of mental health. Gaps in this literature will be discussed and a comparison to research on maladaptive defenses will be made in order to fill some of these gaps. Finally, a brief review of a mental health measure and a short-term intervention that were used in this study are presented, in order to clarify why these measures and treatment protocol were chosen for the present study.

Early History of Psychological Defenses

A brief overview of the earliest theoretical work on defense mechanisms illustrates a number of issues that continue to characterize empirical work on defense. Defining specific defense mechanisms and distinguishing between these mechanisms has been a goal and ongoing problem in both past theoretical work and current empirical work. In addition, a change in the theorized relationship between anxiety and defenses provided a framework for much of the current empirical work on excessive defensiveness.

There has been a good deal of confusion regarding definitions and specificity of defense mechanisms. Much of this confusion might be attributable to the fact that Sigmund Freud used differing terms and definitions until 1926 (Erdelyi, 1990; Holmes, 1990; Vaillant, 1990). S. Freud first used the term "defense" in reference to a situation in which the ego came upon "an idea or feeling which aroused such a distressing affect that the subject decided to forget about it because he had no confidence in his power to resolve the contradiction between the incompatible idea and his ego by means of thought-activity" (1959/1895b, p. 47). It appears as though S. Freud is discussing suppression if this "decision" to forget is conscious, but repression if it is not.[7] Regardless, the term "defense" was later abandoned and he did not mention it again until 1926 (A. Freud, 1936). In the meantime, he used the term "repression" both as a specific defense and as a general term to describe multiple defenses ranging from disavowal to suppression (Vaillant, 1990).

[7] There is some suggestion that Freud was not referring to a conscious decision, though it is unclear – earlier in the same paper Freud discusses splitting of consciousness as an "act of will on the part of the patient" but qualifies that he does not, "of course, mean that the patient intends to bring about a splitting of his consciousness" (1959/1895b, p. 46).

S. Freud admits in an appendix to his 1926 work that he used the term repression in place of defense without clarifying the relationship between them (1959/1926). When he reintroduced the term "defense" in 1926, he stated that a general term was required to refer to any means by which the ego deals with conflict that could lead to neurosis, while he reserved the term "repression" for the special method of defense, common to hysteria in his view, whereby unacceptable thoughts, images, or memories are confined to the unconscious (Laplanche & Pontalis, 1973). In this 1926 work, S. Freud also differentiated several other operations from repression, such as isolation of affect and regression, all of which were to fall under the rubric of defense.

S. Freud's work provided scaffolding for Anna Freud's later work on defenses and, in turn, subsequent empirical work on defensive operations. Two of A. Freud's primary contributions were the further theoretical specification of defense mechanisms and the development of conceptual means by which such mechanisms could be classified. One of the means of classification A. Freud suggested was the use of developmental stages. A. Freud (1936) stated, for example, that if repression is the withholding of an affect or impulse from the conscious ego, it cannot be possible for repression to be in operation before ego and id are differentiated. Similarly, if sublimation involves displacement of an impulse in accordance with what is socially acceptable, one would have to *know* what is socially acceptable in order to implement the defense. Thus, a superego must have developed before sublimation could operate (1936). Although A. Freud later abandoned developmental classification in favor of one that organizes defenses by their source of anxiety or danger, her developmental classification remains the subject of much research on maladaptive defenses today (e.g. Cramer 1988).

Empirical research on maladaptive defenses per se is not the subject of the study at hand. However, such research fills an important gap relevant to the present study on excessive defensiveness by providing a structure for understanding what the excessive defensiveness literature is actually measuring. Research on excessive defensiveness has not elaborated on or specified between more and less mature defenses as maladaptive defensiveness research has attempted to do. Excessive defensiveness research also uses terminology for defenses that is potentially misleading and so had to be clarified for the present effort, as will be discussed below.

A change in S. Freud's theory from repression causing anxiety to anxiety motivating repression made possible nearly all the current literature on excessive defensiveness. S. Freud's use of his earlier term, repression, in 1926 represented a revolutionary change in his theory regarding the origin of anxiety and its relationship to defensive operations. Prior to this, S. Freud had suggested that defenses were constructed against id impulses because acting on such impulses could be dangerous (i.e., result in negative consequences brought on by the external environment such as punishment). The result of "repressing" such impulses was anxiety – a sort of transformation of libidinal energy that had no direct outlet (S. Freud 1959/1895a). S. Freud gave up this theory in 1926 for the alternative that anxiety arises out of fear of trauma. Sterba (1982) stated that this represented a replacement of the older transformational theory of anxiety with a neurotic theory of anxiety whereby anxiety is a danger signal, which motivates the ego to employ defensive operations[8]. Thus, the causal order of operations was reversed from "*repression causes anxiety* to *anxiety motivates repression*" (Sterba, 1983 pg. 78 -- italics

[8] It is noteworthy that this general theory of anxiety as danger signal is also widely accepted in bio-behavioral theories of anxiety (see Gray, 1994).

in original). Much current empirical work on excessive defensiveness is based upon this newer theory of neurotic anxiety – specifically, the idea that defenses serve to keep anxiety and distress out of awareness.

Empirical Work on Excessive Defensiveness

Recall that empirical work using physiological measures suggests that psychological defenses exist and that there is a cost for depending on excessive defenses in terms of impaired mental and physical health, as well as impaired information processing. Furthermore, some research suggests that excessive defensiveness distorts self-report of mental health. The goal of this section is to detail some of the empirical work in these areas as well as address questions concerning this area of research, including the supposition that some degree of defensiveness may be healthy.

Before delving into the empirical literature on excessive defensiveness, an issue articulated by Anna Freud may be appropriate to mention. She theorized that, "all the defensive measures of the ego... ...are carried out silently and invisibly. The most we can ever do is reconstruct them in retrospect" (A. Freud, 1936, p. 8). She further explains that we cannot actually observe defenses per se and that their operation only becomes apparent through the apparent absence of affects that normally would be present (affects that are, for instance, repressed, disavowed, isolated, or suppressed.). Therefore, if she is correct, empirical research on defenses presumably cannot quantify actual excessive defensiveness per se. Nevertheless, multiple findings across several lines of research are consistent with what would be predicted by a theory of defensive operations. Such studies find an absence of self-reported affects where physiological data suggest such affects are present (Dozier & Kobak, 1992; Shedler et al., 1993; Hughes et al., 1994;

Pennebaker, 1997; Weinberger et al., 1979). Such studies also demonstrate physical and mental health consequences associated with denying the presence of uncomfortable affects, where other data (physiological and observational) suggest such affects are present (Jensen, 1987; Pennebaker, 1985; Pennebaker & Susman, 1988; Perry & Cooper, 1989; Shedler et al., 1993; Vaillant & Drake, 1985).

Physiological and behavioral evidence for defensive operations. Physiological data provide one kind of evidence for defensive operations. Several researchers have found that various ways of quantifying an individual's tendency to deny psychopathology correlate with increased heart-rate and blood pressure and increased galvanic skin conductance, consistent with a theory of excessive defensiveness.

Shedler et al. (1993) used the combination of a clinical interview using the Early Memory Test (Mayman, 1968) and self-report neuroticism (Eysenck Neuroticism Scale, Eysenck & Eysenck, 1977) to group participants into the following categories: manifestly distressed (high self-report neuroticism and clinically judged as distressed), healthy (low self-report neuroticism and clinically judged as healthy), and those portraying "illusory mental health" (low self-report neuroticism, but clinically judged as distressed), the latter purportedly as a result of defensive denial of psychological distress. For all of these participants, heart rate and blood pressure were recorded several times while they performed multiple tasks intended to increase anxiety (such as math problems, Thematic Apperception Test [TAT, Murray, 1943], and a phrase association task). Across these tasks, participants with illusory mental health demonstrated greater cardiac reactivity than healthy or even manifestly distressed participants. Shedler et al. (1993)

replicated this finding using multiple means of clinical judgment at well (including trained non-clinicians as clinical judges and a Q-sort procedure).

Weinberger et al. (1979) measured defensiveness differently than Shedler et al. (1993), but obtained similar results with both behavioral and physiological measures. Weinberger et al. hypothesized that repressors (those who deny having heightened levels of anxiety even though they frequently manifest non-verbal behaviors that suggests anxiety), when completing a sentence completion task with neutral, aggressive, and sexual content, would manifest greater skin conductance, heart rate, and forehead muscle tension, as well as longer reaction times, and greater content avoidance in comparison to controls, all interpreted as indicators of distress.

Weinberger et al. used the Taylor Manifest Anxiety Scale (Taylor, 1953) in combination with the Marlowe-Crowne Social Desirability Scale[9] (Crowne & Marlowe, 1964) to distinguish between those who were repressors (low score on the Marlowe-Crowne Social Desirability Scale [i.e. defensive], low score on the Taylor Manifest Anxiety scale), anxious (average score on the Marlowe-Crowne Social Desirability Scale, and high score on the Taylor Manifest Anxiety scale), and non-anxious (average-scores on the Marlowe-Crowne Social Desirability Scale, low scores on the Taylor Manifest Anxiety scale). Their results showed that repressors evidenced greater content avoidance and longer reaction times than non-anxious and anxious individuals, when responding to the sentence completion task. Repressors also demonstrated greater heart rate, sweat gland activity, and forehead muscle tension when responding to the task, demonstrating both behavioral and physiological markers of defensiveness.

More physiological support for defensive operations comes from research using the Denial scale of the MMPI (Little & Fisher, 1958). The Denial scale consists of items querying poor interpersonal relationships, feelings of hostility, and suspicion. Elevated scores on this scale suggest a lack of insight and "anti-intraception" (pg. 306). Denial elevations are commonly found on profiles where the endorsement of pathology appears minimized (Little & Fisher, 1958).

One study using the Denial scale required participants to view a subincision film while heart-rate and blood pressure were monitored (Lazarus & Alfert, 1964). Those who were classified as deniers by the Denial scale reported less distress while watching the film than controls, but evidenced greater distress via heart-rate and skin conductance than controls. In a similar study (Jorgensen, Schreer, Baskin, & Kolodziej, 1992), participants completed the Denial scale and then completed a challenging arithmetic task while their heart-rate was monitored. After the task participants completed questionnaires regarding their anger and anxiety during the task. As expected, high Denial scores predicted low self-reports of anxiety and hostility during the task, but predicted high heart-rate during the same task.

Hughes et al. (1994) provided physiological evidence of defensive operations along a different line of research than those listed above. They hypothesized that inhibiting thoughts and emotions is stressful and does not lead to a resolution of painful experiences. Therefore such inhibition should lead to increased physiological activity. Conversely, they proposed that actively confronting painful emotions and memories might lead to a momentary increase in physiological activity as this requires effort, but a

[9] This scale, in combination with the Taylor Manifest Anxiety scale, is thought to measure affect inhibition/defensiveness rather than socially desirable responding per se (Crowne & Marlowe, 1964;

decrease in physiological activity over the long run. Hughes et al. monitored galvanic skin conductance of participants while they wrote about a traumatic event. Moment-to-moment skin conductance measurements were linked to each word and phrase in participants' essays. Their results were inconsistent for the active confrontation of painful emotions, but defensive strategies were consistently related to increases in electrodermal activity. They report that skin conductance increased when denial and undoing were employed. As with above studies, this research provides further support for defensive operations through physiological measures.

Dozier and Kobak (1992) demonstrate yet another line of research that provides physiological support for defensive operations. They hypothesized that some individuals use a deactivation strategy – a turning of attention away from attachment memories and their related affects -- to cope with the anxiety and painful feeling brought up in an attachment interview. Such individuals report extremely positive relationships with attachment figures and express little emotion, appearing unperturbed, while their recall of attachment memories is poor and they appear to minimize the importance of early attachment experiences. If this type of response did, indeed involve deactivation, Dozier and Kobak believed that this would be evident in skin conductance. As expected, individuals who implemented what Dozier and Kobak thought to be a deactivating strategy, showed marked increases in skin conductance, beyond that of controls, when asked about experiences of separation, rejection, and threat from parents.

The above studies provide findings across several lines of research that are consistent with what would be predicted by a theory of defensive operations. Namely, distress is verbally denied while several behavioral and physiological indicators of such

---

Wiesenthal, 1974). Further information on this scale is detailed in the methods section.

distress are present in individuals who are thought to be defensive. Nevertheless, virtually all of the above studies suffer from a common flaw -- they imply they are measuring a specific defense without a means to differentiate between possible defensives operations at work.

Thus, it is unclear whether denial per se, is responsible for the cardiac responsiveness that Shedler et al. (1993) report. Suppression could just as easily be responsible as denial since we do not know to what degree participants who did not endorse distress were actually unaware of the distress they did not report. In research using the Marlowe-Crowne Social Desirability scale, the use of the term "repressor" is misleading because it implies that repression is the particular defense at work in these individuals (Westen, 1998). While those researching repressive defensiveness define it as affect inhibition/ disavowal (Crowne & Marlowe, 1964; Davis, 1987; Mikulincer & Orbach, 1995; Weinberger et al., 1974), it suffers from the same shortcomings as Shedler et al.'s use of the term "defensive denial." Namely, the nature of the Marlowe-Crowne Social Desirability Scale makes it unclear whether disavowal, repression, inhibition, or suppression is at work in any given individual who responds to the Marlowe-Crowne Social Desirability Scale and the Taylor Manifest Anxiety scale in a "repressive" way. Similarly, researchers using the Denial scale call it either a measure of repression or denial, despite having no means by which they could discriminate between defenses. Hughes et al, 1994 seem to use defenses, suppression, and denial synonymously, failing to define any of these terms in their study.

One could further argue that these studies do not demonstrate the existence of *any* defense. It is possible that those classified as defensive could have known exactly what

affects they were experiencing, but chose not to inform researchers of this via self-report. However, it seems unlikely that this would explain all the indicators of distress, across several studies, that were elevated in "defensive" individuals even beyond that of individuals who admitted their distress – a finding that is, again, consistent with a theory of defensiveness. Regardless, the above findings do not speak to the idea that defenses can be excessive or costly -- an issue that will be discussed below.

Physiological and mental health costs of defensiveness. Evidence suggests that some types of defensiveness are associated with poor psychological and physical health and are thus thought to be immature or excessive. Perry and Cooper (1989) followed individuals with several different psychological diagnoses (borderline personality, antisocial personality, bipolar II, sub- threshold character pathology) over the course of two years, measuring defensiveness, global functioning, and self and observer-reports of symptoms and distress. They found that so-called immature defenses were associated with more symptoms and worse global functioning. Interestingly, defenses falling under the prototype of disavowal were associated with worse observer-rated symptoms but better self-report symptoms, which Perry and Cooper (1989) concluded was a marker for a lack of subjective awareness of symptoms. It is unclear how the authors would know whether participants were unaware of symptoms or chose not to report such symptoms, but their results do support the idea that some (immature defenses) are associated with poor mental health.

Vaillant and Drake (1985) found a relationship between immature defense and DSM-III Axis II diagnosis as well as observer global mental health ratings. Clinicians rated 307 inner-city men on several aspects of psychopathology and resilience. Among

these measures were independent observer ratings of defensiveness, DSM-III axis II diagnoses, and global mental health. As expected, those who were judged as the healthiest used the most theoretically mature defenses (such as altruism, sublimation, and humor). Likewise, those with the most severe psychosocial impairment used the most theoretically immature defenses (dissociation, hypochondriasis, and disavowal).

Pennebaker and several colleagues have generated a great deal of evidence suggesting that excessive defensiveness leads to *physical* health costs. Pennebaker's view is that among those who have experienced psychological trauma, there are a significant number of people who inhibit their thoughts and feelings about the trauma, hiding it from others and, often, from themselves (Pennebaker & Susman, 1988). This inhibition, according to Pennebaker, is a long-term physical stressor, which leads to adverse health effects, and increased ruminations about the trauma. Pennebaker believes that having people disclose or confront previous trauma will improve health and reduce rumination about the trauma (Pennebaker & Susman, 1988).

Pennebaker and Susman (1988) surveyed 200 participants about recent and past trauma and how much they confided in others about the trauma. They also assessed participants' major and minor health problems (such as cancer, heart disease, headaches and colds). Those with recent or past trauma, who did not confide in others, were more likely to have had major and minor health problems in the last year, than those with no trauma or those with trauma who did discuss the experience with others. In a multiple regression analysis, early trauma significantly predicted health problems, and whether or not participant's confided in others about the trauma added additional predictive power (Pennebaker & Susman, 1988). In another study, Pennebaker and O'Heeron (1984)

assessed widows the year following their spouses' suicide or death by car accident. Illness rates and the degree to which they confided about their loss to others were assessed. Similar to Pennebaker's previous study, the more participants confided in others about their spouses' death the fewer health problems they had. While the above studies yield some compelling evidence that not disclosing thoughts and feelings about trauma can be related to heath problems, it may not speak to defensiveness: failing to talk about painful feelings with others is not necessarily the same as failing to think about or acknowledge such feelings at all. Furthermore, results may be confounded by quality of social support. Those who chose to confide may have had better social support than those who did not. This quality of social support could have been responsible for the better health of the confiding group.

Pennebaker and colleagues attempted to further investigate the degree to which thinking about and processing trauma, itself, might have a causal relationship with physical health. Pennebaker and Beall (1986) asked participants to write about a traumatic event for 15 minutes each day. Control participants were asked to write about something trivial for the same amount of time. The experimenters collected data on health center visits for all participants 3 months before and 6 months after the experiment. Those who wrote about their thoughts and feelings surrounding a traumatic event evidenced fewer illness visits to the health center 6 months post experiment. While it is unknown to what degree control participants were inhibiting thoughts about trauma, actively thinking painful thoughts and feelings surrounding a traumatic event while journaling clearly lead to better physical health.

In a similar study Pennebaker, Glaser, & Glaser (1988) measured immune system functioning in two groups of college students. Again, one group wrote about painful, traumatic experiences and another wrote about trivial issues, this time for 20 minutes each day. Experimenters tested immune functioning one day before, one day after, and six weeks after the experiment. Participants who wrote about thoughts and feelings surrounding a traumatic experience demonstrated significant improvement in immune functioning over the course of the study and differences were maintained 6 weeks after the study as well. Overall, Pennebaker's research does not directly link disclosure (or lack there of) with his theorized consequences of inhibition (and therefore, consequences of excessive defensiveness). In his journaling studies, it is also feasible that some differences between groups were the result of the control group being irritated with being required to write about trivial topics. However, his work as a whole strongly suggests that actively thinking about painful thoughts and feelings while taking positive action (journaling/confiding) leads to greater physical health.

Shedler et al.'s (1993), work also lends support to the theory that there are physical costs for depending on defenses. You will recall that participants in this study portraying illusory mental health (low self-report neuroticism, but clinically judged as distressed), demonstrated greater cardiac reactivity than healthy or manifestly distressed participants. The illusory mental health group (i.e. defensive deniers) manifested such large spikes in heart rate and blood pressure that they were at a level that has been found to correlate with damage to the inner lining of coronary arteries, and put people at risk for arteriosclerosis and heart disease (Krants & Manuck, 1984; Manuck, Kaplan & Clarkson, 1983).

Converging evidence for the physical health costs of excessive defensiveness comes from Jensen's (1987) work on psychological factors that influence the course of breast cancer. As with the Weinberger et al. (1979) study above, Jensen used the Marlowe-Crowne Social Desirability Scale and the Taylor Manifest Anxiety scale to classify participants as repressors (low score on the Marlowe-Crowne Social Desirability Scale [i.e. defensive], low score on the Taylor Manifest Anxiety scale), high-anxious (average score on the Marlowe-Crowne Social Desirability Scale, and high score on the Taylor Manifest Anxiety scale), and low-anxious (average-scores on the Marlowe-Crowne Social Desirability Scale, low scores on the Taylor Manifest Anxiety scale). Jensen found that despite reporting few negative feelings, repressors' records demonstrated (1) shorter periods of remission following initial diagnosis and treatment, (2) greater cancer spread by follow-up, (3) greater medical deterioration by follow-up, and (4) higher death rates by follow-up. Repressors' overall outcomes were worse than both low-anxious and high-anxious groups. Jensen also tested the possibility that biological variables might have influenced behavioral data (i.e. worse prognosis at the outset leading to increased repressive coping). She found that prognosis at the time of diagnosis, disease stage at the time of diagnosis, time lapse since diagnosis, age of onset, disease course, and medical status at the onset of the study did not account for her psychological findings. This strongly suggests that repressive coping exacerbated the disease process in these cancer patients.

The preceding studies have found psychological and physical health costs for excessive defensiveness, using various methods to observe this phenomenon. However, like other studies on excessive defensiveness, definitions are misleading – Jensen (1987)

and Schedler et al. (1993) may be measuring anything along a continuum from disavowal to suppression and Pennebaker admits that his work may not be tapping inhibition at all, even though inhibition is an important piece of his theoretical construct. Nevertheless, the convergence of this evidence across various methods of measuring *excessive* defensiveness (regardless of which defenses, in particular, are measured) is compelling.

The costs of conscious thought suppression. It should be added that there is substantial evidence that thought suppression has a psychological cost, even though it is considered a more adaptive defense (Cramer, 1989; Holmes, 1990; Vaillant, 1990). Gross has conducted several studies on the effects of affective suppression, that is, inhibiting the expression of an emotion that is consciously recognized. In one such study, Gross and Levenson (1997) had participants view films with neutral, amusing, and sad affective valences. The experimental group was told not to give any physical signs of what they were feeling. The control group was told simply to view the films. Gross and his colleagues found that those told to inhibit the expression of their affect demonstrated physiological signs of sympathetic activation, beyond that of the control group, during the amusing and sad films, but not during the neutral toned film. Gross and Levenson (1997) suggest that this increase in sympathetic activation may lead to impaired sensory intake of information from ones environment, leading to diminished cognitive performance.

Richards and Gross (2000, study 1) tested the possibility that affective suppression would decrease cognitive performance by showing participants a film known to induce negative affects characterized by sadness, anxiety, and anger. As with the previous study, the experimental group was told not to give any physical signs of what

they were feeling, while the control group was simply told to watch the film. As expected, affective suppression lead to poorer memory for the film's auditory and visual details.

In a second study, Richards and Gross (2000, study 2) tested whether their findings extend to situations outside the laboratory. They used questionnaires to assess expressive suppression, reappraisal (a more adaptive method for coping with painful or anxiety provoking affects), and participants' own perceptions of how well they remembered situations in which they used these strategies. Richards and Gross (2000, study 2) also required participants to recall situations in which they used either affective suppression or reappraisal with a daily diary. Finally, participants were tested weekly for recollection of what they wrote in their daily diaries, as a more objective measure of memory. As they expected, Richards and Gross (2000, study 2) found that the use of affective suppression was associated with poorer self-report and objective memory measures.

One empirical issue with Richards and Gross' (2000) work is that we have no way of knowing what information may have been forgotten before participants wrote in their daily diary. Richards and Gross (2000) theorize that suppression inhibits the taking in of information, but once information has been written in a diary it has already *been* taken in. In this way, it would seem to be memory retrieval rather than processing that has been tested. Nevertheless, when taken as a whole, Richard and Gross' (2000) work suggests that there are costs, even for conscious thought suppression. This is important to note for the present study because of the implication that it may not matter exactly which defense keeps a person from endorsing psychopathology on self-report questionnaires. Even

suppression may have a mental health cost in terms of using available information to make choices. That is, if suppression causes impaired sensory intake (or impaired recall of information) then not all information will be available to make constructive choices in a distressing situation.

Possible distorting effects of excessive defensiveness on self-report. If it is true that some people tend to excessively deny their experience of painful or anxious feelings, it follows that they may do so across self-report measures of mental health. In support of this possibility, Davidson and MacGregor (1995, cited in Westen, 1998) found that defensiveness moderates the discrepancy between self-reports and informant reports of an individual's neuroticism. In individuals who displayed moderate or little defensiveness, self-reports and informant reports of neuroticism correlated with each other modestly. However, in individuals who were highly defensive, self-reports and informant-reports of neuroticism correlated strongly in the negative direction. That is, in individuals who were highly defensive, the less they rated themselves as neurotic, the more an observer-rated them as neurotic.

Shedler et al. (1993) had similar findings when they studied this issue from a physiological perspective. Among participants who were judged to be distressed by a clinician, the lower the self-report of neuroticism, the higher the level of cardiac reactivity, supporting the hypothesis that the less anxious the self-portrayal of a distressed person (or the more defensive the individual was), the more distressed the individual became (Shedler et al).

Much of the research above detailing physiological and psychological costs of defenses suggests that excessive defensiveness distorts self-report of psychopathology.

Across the repressive coping literature, exceedingly few individuals obtain high scores on both the Marlowe-Crowne Social Desirability Scale (a measure of excessive defensiveness, as seen above) and the Taylor Manifest Anxiety scale – a self-report measure of anxiety (Davis, 1985; Davis & Schwartz, 1987; Mikulincer & Orbach, 1995; Weinberger et al., 1979). This suggests that those who are defensive tend not to admit to anxiety.

Some may argue against the distorting effects of defenses on self-report mental health measures by citing the literature on relevant outcomes predicted by such self-report measures (see Block, 1965; Watson & Clark,1984). However, the present study does not argue simply that self-report measures of mental health are inaccurate. As Shedler et al. (1993) contend, self-report measures of mental health are likely accurate when individuals admit to poor psychological health,[10] but increasingly ambiguous at higher levels of self-reported health. This would probably not have the effect of eliminating relationships between self-report mental health measures and relevant outcomes but it would probably attenuate such relationships. Furthermore, the validity of such measures is typically assessed by comparing them to other self-report measures (see Beck, Ward, Mendelson, Mock & Erbaugh, 1961; Spielberger, 1983). As Shedler et al. (1993) point out, comparing one invalid measure to another invalid measure and finding they get similar results does not validate either measure.

An alternative hypothesis: "Ignorance is bliss." Some have argued that it is both common and psychologically healthy to believe one is better adjusted than one actually is. Taylor and Brown (1988) conducted one of the most influential surveys on this

[10] Of course, malingering or a "cry for help" may bias responses on this end of the continuum as well, but such questions are beyond the scope of the present study.

literature of self-enhancement, suggesting that people who are psychologically healthy consistently have an unrealistically positive perception of themselves and harbor an exaggerated belief in their own control and mastery. This seems to suggest that being "well defended" may promote psychological health, at least up to a point.

The view that psychologically healthy people consistently self-enhance has come under criticism for several reasons. One of the critiques most relevant to the present study is that self-reports of adjustment are typically used to identify the characteristics of self-enhancing people. Thus, it is not surprising that people who self-enhance also describe themselves as being well adjusted psychologically and endorse few symptoms of psychopathology (Colvin, Block, & Funder, 1995). Thus, the present author interprets findings of studies that tout the psychological benefits of self-enhancement as potentially a symptom of the distorting effects of defensiveness on self-report measures, rather than evidence of the healthy benefits of self-enhancement.

This flaw in self-enhancement research is supported by studies on self-enhancement that use observer reports of mental health rather than self-reports. Colvin et al. (1995) found that self-enhancers, when rated by doctoral level psychologists and advanced graduate students with a California Q-sort procedure, were rated as guileful, deceitful, having brittle ego defenses, and thin skinned. On the other hand, non self-enhancers were described as introspective, complex, interesting, and intelligent people.

John and Robins (1994) required MBA students to complete a task in small groups and rate their own as well as others' contributions to completing the task. Whereas most individuals self-enhanced to some degree (60% rated themselves one point higher than observing psychologists), those who self-enhanced to a slightly larger degree

(rated their contribution 2 or more points higher) were judged by clinicians and peers as narcissistic on several measures, suggesting poorer mental health.

In short, a slightly self-enhancing bias may be typical, but more extreme self-enhancement is likely psychologically unhealthy. Some theorize that this self-enhancement leads to loneliness due to subsequent rejection by peers and reinforces a narcissistic defense against this rejection (Colvin et al., 1995). It is likely that extreme self-enhancement is related to excessive defensiveness. Thus, utilizing extreme self-enhancement might mean inheriting the psychological and physical health costs that go with such excessive defensiveness.

The theory that minor self-enhancement may be normative brings up a possible nuance in the relationship between defensiveness and mental health. While several studies mentioned above demonstrate negative consequences for excessive defensiveness, a normal degree of defensiveness is thought to protect a person from consciously experiencing excessive anxiety or guilt (Westen, 1998). Thus, defensiveness most likely has a curvilinear relationship with mental and physical health. Being excessively defensive is probably unhealthy, but not being defended enough is also likely unhealthy. This was taken into account in the present study by means of analyses designed to tap curvilinear relations, as described later.

Empirical Work on Adaptiveness of Defenses

While the present study did not analyze defenses in terms of adaptiveness or maturity, there are a number of reasons it is important to outline past work in this area. Research looking at defenses in terms of adaptiveness/maturity has several limitations. However, it does provide a means by which to understand what research on excessive

defensiveness might actually be tapping, because the latter has defined defenses rather loosely, where it defines them at all. Also, covering measures of the adaptiveness/maturity of defenses affords an opportunity to elaborate why such measures may not be as useful for the present study.

Most researchers still agree with S. Freud's supposition (1959/1895b) that the broad role of defense is to protect a person from consciously experiencing excessive anxiety or guilt (Westen, 1998). Actually defining or quantifying specific defenses, however, has been difficult. There has been little agreement as to how many defenses there are and how to reliably distinguish between specific defenses such as disavowal and repression (Cooper, 1992; Davis & MacGregor, 1998; Vaillant, 1971, 1990), and some have argued that distinct defense mechanisms cannot be delineated at all (Brenner, 1982). Several researchers have tried to delineate different defenses and classify them in terms of adaptiveness or maturity. These attempts have taken several forms.

Observer-rated measures. Some researchers have attempted to arrange defenses into a hierarchy by means of observer-rated measures. This practice was preceded by Anna Freud's hierarchy of defenses (A. Freud, 1936) and continued with the empirical work of Haan (1963) and Vaillant (1971), using clinical ratings of defenses. Vaillant's group later used a Q-sort methodology to quantify various defenses in terms of a hierarchy of adaptiveness (Vaillant, 1990; Roston, Lee, & Vaillant, 1992). Vaillant's Q-sort, the most well know Q-sort measure of defenses, consists of 15 defenses across 3 broad categories: mature defenses (including defenses such as sublimation, humor, and altruism); intermediate/neurotic defenses (including displacement, isolation, and reaction formation); and immature defenses (including projection, dissociation [neurotic denial],

and hypochondriasis). Vaillant and colleagues have used this Q-sort to rate interviews with over 500 high achieving men followed from their late teens/early twenties into their late fifties (participants were interviewed for two hours approximately once every 10 years). Reliability for some defenses was poor. For example, the interrater correlation for dissociation (neurotic denial) was r=.01 (Vaillant, 1990) but, regardless of rater, frequent use of this defense correlated with poor outcome (Valliant, 1990). Reaction formation, also an "immature" defense, correlated with poorer outcome and repression correlated with low ego development (Vaillant, 1990). Finally, the overall maturity level of defenses used was correlated with global functioning (Perry & Ianni, 1998).

Overall, Vaillant's Q-sort appears to be a valid and reliable measure (except for a handful of specific defenses with poor interrater reliabilities) with a rich, longitudinal data set to support it, albeit, on a limited sample. While no data reported by Valliant's group suggest that the defenses actually belong in the specific "maturity" categories that they are classified in, there are data to suggest that they do fit along a continuum of adaptiveness that parallels the hypothesized continuum of maturity (Vaillant, 1990). His measure falls short in terms of its usefulness for the present study due to the lack of reliability on denial, an important defense to be able to quantify and, most practically, its unwieldiness in terms of time required to administer the measure – 2 hours.

Self-report measures. Other researchers have used a questionnaire format to group defenses by adaptiveness or maturity. The most widely used of such measures is the Defense Mechanism Inventory (DMI, Gleser & Ihilevich, 1969; Davidson & MacGregor, 1998). The DMI consists of 10 stories describing conflicts that purportedly tap 5 defense clusters. These clusters include Turning Against the Self (internalization of

external problems and inhibiting aggression), Turning Against the Object (displacement and identifying with the aggressor), Projection (placing one's own unacceptable feelings on someone else), Reversal (negation, denial, reaction formation, and repression), and Principalization (intellectualization, isolation, and rationalization). Tests of the validity of this scale have had mixed results (Cramer, 1988). One strength of the measure is that the Reversal scale, ostensibly a measure of immature defenses, appears to correlate with criterion measures better than other scales do. Reversal has been consistently related to denial, primitive defenses, repression, and avoidance, as well as a tendency to overvalue one's performance and greater psychopathology. A weakness of the DMI is that Principalization correlates with much of the same defenses as the Reversal scale, even though it is supposed to tap more adaptive/mature defenses (Cramer, 1988). It has also been argued that while the concepts behind Projection and Turning against the Object are sound, the items making up these scales don't tap defensiveness at all (Davidson & MacGregor, 1998).

The DMI has several issues that make it inappropriate for use in the present study. As noted above, the Reversal and Principalization scales correlate with the same variables despite being, theoretically, on opposite sides of the continuum in terms of adaptiveness. Also, usage of the DMI and other self-report defense questionnaires is rife with the theoretical dilemmas involving self-report (Davidson & MacGregor, 1998) that the present study seeks to investigate. Most importantly, the DMI (and other self-report defense questionnaires such as the Defensive Style Questionnaire [Bond, GarDenialer, Christian, & Sigal, 1983]) assumes that people will accurately report on their own defensive style, which, theoretically, should be unconscious, at least much of the time

(Davidson & MacGregor, 1998; Westen, 1998). Therefore, such questionnaires appear susceptible to distortion by virtue of the very defenses they are supposed to measure and do not provided a vantage point from which to evaluate self-report.

What the adaptiveness literature tells us about excessive defensiveness. Although measures of adaptiveness of defenses will not be used in the present study, they may fill a theoretical gap that is present in the literature on excessive defensiveness. The literature on excessive defensiveness, as a whole, poorly defines what defenses it is measuring. For example, it has been argued that Shedler et al.'s study (1993) on defensive denial does not measure denial alone (Westen, 1998) and that the repressive coping literature does not measure repression per se (Holmes, 1990). Furthermore, researchers do not agree on what defenses these studies have actually accounted for. Vaillant states that the repressive coping literature measures dissociation (also known as neurotic denial in his Q-sort measure) and isolation of affect (Vaillant, 1990). Weinberger et al. (1979) contend that repressive coping constitutes an inhibition of affect, and Westen (1998) states that the literature on repressive coping actually measures "a tendency to avoid feeling emotions as a way of managing distress more akin to isolation of affect" (p 27). These various objections point to an agreement between researchers, however, that whatever is measured in the excessive defensiveness literature, by and large these defenses, at least theoretically, fall into an immature/maladaptive defense category across scales. On Vaillant's Q-sort they would be categorized as immature and on the DMI they

would fall in either the Reversal or Principalization category, each of which correlates with more primitive defenses and greater psychopathology.[11]

In summary, it appears we cannot know exactly what is being measured in the excessive defensiveness literature because we cannot see exactly *why* "repressors" or "defensive deniers" or those using "deactivating strategies" report a lack of distress despite physiological and behavioral data that suggest otherwise. There is agreement across researchers that excessive defensiveness literature measures defenses, which fall within a "maladaptive" range.

It should be noted that for the purposes of this study, the specific defenses contributing to inaccurate self-reports of mental health do not matter. Whether it be denial/disavowal, repression, isolation of affect or suppression, this response style has been associated with physiological and psychological costs in terms of health, across several studies. Even conscious thought suppression has cognitive and physiological costs (Gross & Levenson, 1997; Richards & Gross, 2000). These costs make it especially important to take defenses into account when measuring the effectiveness of an intervention.

Gender and defensiveness. Few published studies on relationships between gender and defensiveness exist. However, apparent differences between men's and women's self-report of psychopathology, along with the few studies that exist on gender and defensive functioning, demonstrate a trend toward differences in ways men and women use psychological defenses. There are consistent findings that women report more anxiety and depression than men (Nolen-Hoeksema, 1989). This may indicate that

[11] Recall that even though Principalization was thought to be more adaptive by Gleser & Ihilevich (1969) it correlates with most of the same criterion measures as the Reversal scale -- denial, primitive defenses,

women are more prone to psychopathology, but may also indicate that men are more likely to defensively deny psychopathology that is present. In support of the latter explanation, men have been found more likely to score 0 on the BDI than women (Joiner, Schmidt, Lerew, et al., 2000). As stated above, such scores may be more indicative of defensive denial of pathology than they are of a total absence of depressive symptoms (Kendal et al., 1987). Furthermore, defensiveness has been related to low self-reports of depression more so in men than in women (Joiner et al, 2000) and men have found to generally utilize more immature defenses (Bullitt & Farber, 2002). However, it is clear that defensiveness is related to low self-report of depression in women as well (Scafidi et al., 1999). Such potential gender differences in terms of defensiveness will be taken into account in the present study.

Results from the above gender-related defensiveness studies are difficult to interpret because across these studies low self-report pathology scores coupled with high scores on defensiveness measures are assumed to indicate inaccurate self-report of mental health. These studies do not utilize a second measure of mental health to verify that the self-reports of mental heath are, in fact, less accurate in excessively defensive individuals. This highlights the need for a second measure of mental health that is not as susceptible to the distorting influence of defensiveness in the present study.

Characteristics of a "More Accurate" Mental Health Measure

As stated above, one goal of this study is to determine if excessive defensiveness will cause an individual to deny pathology on self-report questionnaires that is actually present. However, it cannot be assumed that an individual is not mentally healthy simply because he or she denies pathology and appears excessively defensive on a measure

---

repression, and avoidance, as well as greater psychopathology.

thereof. A non-self-report measure of mental health is needed in this study to obtain evidence as to whether those who are excessively defensive and report a lack of psychopathology are, indeed, not as psychologically resilient as they report.

Mental health can be viewed as a broad, 'umbrella' term covering multiple and overlapping constructs. Some of these constructs could include level of psychopathology, global functioning, and ego functioning. Generally, self-report measures of mental health used in studies detailed earlier tap the first (level of psychopathology) construct. Several observer-rated mental health measures tap either level of psychopathology or global functioning constructs. Although observer report scales such as the Hamilton Rating Scale (Hamilton, 1960) or the Global Assessment of Functioning (Endicott et al., 1976) might be useful in terms of assessing these aspects of mental health, many would argue that such scales may still be susceptible to the denial of psychopathology. If an individual is able to keep distress out of his or her own awareness, it follows that the same individual may be able to hide outward signs of this distress, at least to some degree, as well (Westen, 1998). Thus, a useful criterion measure of mental health would also have to incorporate information beyond an individual's initial presentation.

In addition to observer ratings, a projective approach may make a criterion measure of mental health less susceptible to distorting effects of defensiveness. The projective hypothesis, an assumption underlying all projective measures, suggests that individuals reveal aspects of their own personality in the process of making sense of ambiguous stimuli (Frank, 1948). For instance, individuals who create stories in response to TAT cards (Murray, 1943) are assumed to project enduring expectations,

fears, and desires that may be influential in their daily interactions, but unavailable to conscious introspection (see Conklin & Westen, 2002). Having observers rate such projections may be a means of circumventing defenses that could distort self-report of mental health status.

Some researchers in the past have attempted to incorporate observer ratings and/or projective data into their mental health measures. Such measures have included clinician ratings (Shedler et al., 1993), the Adult Attachment Interview (Hesse, 1999) or aspects of Exner's Rorschach scoring system (Exner, 1993). However, such measures have suffered from issues such as a lack of reliability (Lilienfeld, Wood, & Garb, 2000) or cost in terms of the time necessary to administer or score the measures.

The Social Cognitions and Object Relations Scale (SCORS, Westen, 1991) may be a reasonable alternative measure of aspects of mental health. The SCORS scale is a means by which an observer (coder) can quantitatively rate narrative data (such as early memories or TAT responses) based on object-relations and social cognitions theories. Additionally, the scale incorporates a psychodynamic perspective, measuring mental health not simply as a lack of psychopathology, but as the *presence* of various theoretically important capacities such as the ability to experience interrelatedness, intimacy, and a wide, appropriate range of affects (Shedler et al., 1994). The SCORS has five dimensions: Complexity of Representations (ranging from poor differentiation between people and egocentrism to rich and psychologically minded descriptions of peoples' personalities), Affect-Tone of Relationship Paradigms (ranging from strongly malevolent to positive and enriching), Emotional Investment in Relationships (ranging from strictly need gratifying, tumultuous, and interchangeable to committed relationships

with mutual sharing and interdependence), Emotional Investment in Values and Moral Standards (ranging from antisocial and aggressive depictions [without a sense that such actions are wrong] to a commitment to both moral ideals and real people), and Understanding of Social Causality (ranging from illogical or distorted narratives to highly coherent narratives).

In terms of overlapping constructs that make up "mental health," the SCORS can be best understood as a measure of ego development. Several theoretical perspectives on ego development are tapped by the scores subscales cited above. For instance, Loevinger (1976) described part of ego development as entailing increasing differentiation among one's perceptions of self, social world, and their interrelationship. In her model, individuals develop from stages characterized by impulsivity, fearfulness, stereotyped attributions about other's thoughts and feelings, and dependant and exploitative relationships to stages characterized by emotional interdependence, an appreciation for differences, more complex attributions about others thoughts and feelings, and relationships characterized by mutuality and empathy. Similarly, Kegan (1982) describes stages of development that lead toward increasing differentiation between oneself and the world and increasingly complex understandings of the others that result in being 'in relation' to others rather than 'enmeshed with' them. These perspectives on ego development are captured well in the Complexity, Affect Tone, Relationships, and Social Causality subscales of the SCORS. Ego development theories that focus on moral development (Kohlberg, 1984; Perry, 1968) have described stages progressing from egocentrism (moral behavior determined by whether or not punishment ensues) to conforming to social rules in order to please other and/or win praise, to 'post

conventional' morals that involve the internalization of principals that take into consideration the needs of others. These theories are accounted for in the Moral Standards subscale of the SCORS.

Although ego development is not a direct measure of psychopathology, the construct is clearly related to it (Browning, 1986; Naom, 1998; Naom & Dill, 1991; Naom et al., 1984; see Hauser & Safyer, 1995 for a review) and the SCORS has been directly related to psychopathology (see below). Thus, although no measure taps all constructs subsumed by the term "mental health," as a measure of ego development the SCORS provided a theoretically useful contrast to self-report measures of psychopathology for present purposes.

Summary and Hypotheses

This study examined two questions. First, it examined whether psychological defenses are related to the denial of psychopathology on self-report mental health/psychopathology measures. If so, the validity of such self-report measures is brought into question, because one's own report of mental health cannot be taken at face value. Second, it examined whether defensiveness attenuates the results of a treatment outcome study that utilizes such self-report measures. This attenuation on group means might occur because some individuals who benefited from the intervention could have endorse more psychopathology *after* an intervention than *before* the intervention, if the intervention served to decrease defensiveness in those who were excessively defensive.

Hypothesis1: Self-reported mental health would predict observer-rated mental health at low levels of defensiveness; at higher levels of defensiveness, the relationship between observer-rated mental health and self-report mental health would be weaker.

Prediction: Defensiveness would moderate the relationship between self- and observer-report of mental health.

Hypothesis 2a: After a psychotherapeutic intervention, individuals in the treatment group would show improved observer-rated but not self-reported mental health as compared to controls. Prediction: There would be an interaction on observer-rated mental health between groups and time (pre and post intervention). Observer-rated mental health would be greater at time 2 than at time 1 only for the treatment group. The same interaction between group and self-report mental health would not be significant.

Hypothesis 2b: Defensiveness would exert a suppressor effect on self-report mental health change in the treatment group. In other words, defensiveness would be associated with higher self-reports of mental health. Prediction: Covarying defensiveness, which was expected to be lower at time 2 than at time 1, would make visible an improvement in self-report of mental health at time 2 over time 1 in the treatment group. Such improvement in self-report mental health would not be apparent before defensiveness was covaried (as was predicted in hypothesis 2a).

Post hoc analyses: Several post hoc analyses were performed. One involved splitting participants into high and low defensiveness groups in order to determine relationships between self-report and observer-rated measures in these groups. Other post hoc analyses were performed on a typological configuration of SCORS scales. The n of the SCORS groups created in this analysis was low, and results based on this typological approach were exploratory.

METHOD

Participants

Experimental participants included 31 students enrolled in Psychology 325 at Michigan State University – a course entitled "Affect and Self-Esteem," which is considered an analog of psychotherapy (Kaufman, 1996). This course was the intervention in the present study and its theoretical basis and implementation is described in more detail below. The control group consisted of 45 Michigan State University students enrolled in Psychology 424 (Child and Family Psychopathology). Participants were recruited from each course on the first day of that course. They were told that they were invited to participate in a study examining how the course affects several variables related to psychological adjustment and that participation would involve completing standardized questionnaires about their thoughts, behaviors, and psychological health, as well as completing 6 short stories (spending 3 minutes on each). They were further told that if they participated in the study on that day and again on the final day of the course they would be entered in a lottery to win $100. The participation rate for the experimental group and control group was 47% and 41%, respectively. Demographics of the total sample are displayed in Table 1.

More details on the sample, including demographic comparison of the treatment and control groups, are provided in the results section. More details on the procedure are provided below after a detailing of the measures used.

Table 1

Sample Demographics

Total N	76
Age (M/SD)	21.8/2.5
Grade (M/SD)	3.8/.38
Sex (Male/Female)	16/60
Caucasian (Frequency)	63
African-American (Frequency)	3
Hispanic/Latino (Frequency)	8
Asian (Frequency)	2

Defensiveness Measures

Marlowe-Crowne Social Desirability scale (MCSD): Crowne and Marlowe (1964) developed their 33 item scale because the Edwards (1957), a widely used scale at the time, confounded social desirability with psychopathology. That is, denying an item on the scale could either indicate a desire to appear socially desirable or a true lack of psychopathology (i.e., one could deny the item "I have diarrhea... once a month or more" either to appear normal or because they actually lack the symptom). Crowne and Marlowe (1964) eliminated this problem by only including items that were either socially desirable and occur extremely rarely (e.g. "No matter who I'm talking to, I'm always a good listener) or socially undesirable, but occur frequently (e.g. "I sometimes try to get even rather than forgive and forget"). Later, it became clear to Crowne and Marlowe (1964) as well as others (Arkin & Lake 1983; Nordholm 1974; Mc Crae & Costa1983; Wiesenthal 1974) that rather than identifying those who distort their responses to look good, the scale was identifying an individual difference variable. Weinberger (1992) agreed with Crowne and Marlowe (1964) that the scale actually measured a defensive style whereby an idealized version of the self was maintained (Crowne & Marlowe, 1964), citing evidence that persons who knew high scorers indicated that high scorers

actually attempted to conform to the rigid standards they endorsed on the scale (Strahan & Strahan, 1972).

Alpha coefficients for the scale have been measured at .73 to .88 across several studies (Crowne & Marlowe, 1964; Paulhaus, 1984; Tanaka-Matsumi & Kameoka, 1986; Fisher, 1967). Crowne and Marlowe (1964) have reported a test-retest correlation over one month at .88. The alpha coefficient for this scale in the present sample was .81, and test-retest reliability over 15 weeks was .82. Both were considered acceptable for the present study.

MMPI Denial scale (MMPI-Dn): The development of the 26 item Denial scale came out of the finding that the MMPI Hysteria scale is commonly elevated in medical patients without psychiatric diagnoses (Little & Fisher, 1958). As a result of this finding, Little and Fisher (1958) factor analyzed the scale for this population and found two primary dimensions. The first, which they called the Admission scale, consisted of endorsed physiological symptoms. The second included denied items querying poor interpersonal relationships, feelings of hostility, and suspicion (e.g. "I think most people would lie to get ahead" and "I often wonder what hidden reason another person might have for doing something nice for me"). This was labeled the MMPI-Dn scale. Little and Fisher (1958) state that in "clinical application" elevated scores on this scale suggest a lack of insight and "anti-intraception" (pg. 306). They state that an MMPI-Dn scale elevation is commonly found in profiles where the endorsement of pathology appears to be minimized. MMPI-Dn scale Kuder-Richardson reliability was measured at .75 by Little and Fisher (1958). The alpha coefficient for the scale in the present study was .73

and test-retest reliability over 15 weeks was measured at .72. Both were considered acceptable for the present study.

Self-Report Mental Health Measures

State-Trait Anxiety Inventory (STAI): Spielberger et al. (1983) designed this scale to measure present and general levels of anxiety. The scale consists of 40 items (20 assess trait anxiety and 20 assess state anxiety). The trait scale alpha coefficients were .90 for males (n=324) and .91 for females (n=531) in a college population (Spielberger et al., 1983). The state scale alpha coefficients were .91 for males and .93 for females in the same population (Spielberger et al., 1983). Test-retest reliability of the trait scale ranged from .84 for 1 hour to .73 for 104 days for men and from .76 for 1 hour to .77 for 104 days for women (Spielberger et al., 1983). Validity of the STAI was obtained by comparing controls' scores with psychiatric patients' scores for whom anxiety was a major symptom. Spielberger et al. (1983) report that the scores on the STAI discriminate well between the two groups. The STAI also correlated well with other self-report measures of anxiety such as the IPAT Anxiety Scale (Cattell & Scheier, 1963, cited in Speilberger. 1983) and the Taylor Manifest Anxiety Scale (Taylor Manifest Anxiety scale; Taylor, 1953), with correlations between the three ranging from .73 to .85. The alpha coefficients for the State and Trait scales in the present sample were .92 and .93 respectively. Test-retest correlations for the State and Trait scales in this sample were .56 and .71, respectively. These were in line with expectations and suggested the scales were suitable for use (i.e., one would expect state anxiety to have a low test-retest score).

A literature search yielded 286 studies utilizing the STAI including 29 efficacy and effectiveness studies over the past 10 years. It served the present purposes well because it appeared to be a widely used self-report measure of psychopathology.

Beck Depression Inventory (BDI): Beck, Steer, and Brown (1996) tout the BDI as the "most widely accepted instrument for assessing the severity of depression in diagnosed patients and for detecting possible depression in normal populations" (pg. 1). The self-report measure contains 21 items and has been normed for adolescents 13 years old and older as well as adults (Beck et al., 1996).

Internal consistency of the BDI was measured at .93 for a group of 120 college students (Beck et al., 1996). Test-retest correlation at a one-week interval was measured at .93 as well (Beck et al., 1996). In the present sample, the alpha coefficient for the BDI was measured at .90 and test-retest reliability over 15 weeks was .79. Both were considered acceptable for the present study.

Convergent validity has been established between the BDI and the Beck Hopelessness Scale (Beck & Steer, 1988), the Scale for Suicidal Ideation (Beck, Kovacs, & Weissman, 1979), and the Hamilton Rating Scale for Depression (Hamilton, 1960) at r = .67, r = .37, and r = .71 respectively (Beck et al., 1996).

With regard to ethnic differences, Beck et al. (1996) report that correlations between race and BDI scores were non-significant (comparing whites to non-nonwhites). However, differences were detected between males' and females' overall scores with females reporting a higher degree of depressive symptoms than males (female raw score = 23.61 [SD = 12.31]. male raw score = 20.44 [SD = 13.28]).

Although Beck et al. (1996) caution against using the BDI as a sole diagnostic measure in the test manual, they do so only on the basis that depression occurs conjointly with several other psychiatric disorders. They do not address whether the BDI might be susceptible to the effects of disavowal, suppression, affect inhibition, or repressive coping.

A literature search yielded 825 studies utilizing the BDI including 93 efficacy and effectiveness studies over the past 10 years. Its wide-spread usage suggested it was a good choice of a commonly used self report psychopathology measure to evaluate in the present study.

Observer-Rated Mental Health Measure

SCORS-TAT: The Social Cognitions and Object Relations Scale (SCORS; Westen, 1991) was designed to assess narrative data for implicit aspects of personality functioning from an object-relations and social cognitions perspective (Westen, 1991). Its format is such that administration and scoring are less time consuming than with some other measures with similar qualities. It has been judged to be more valid and reliable than several other projective measures (Lilienfeld et al., 2000) and its projective nature, as well as its observer based scoring system, may make it less susceptible to distortion by defensive responding than self-report measures. The five subscales of the SCORS are intended to assess capacities and developmental achievements that are integral to mental health from an object-relations perspective (Westen, 1991). The Complexity of Representations of People scale is based on the theory that as an individual psychologically matures, his or her representations of people become better distinguished, multidimensional, and integrated. Undifferentiated representations are

scored low on this scale whereas rich representations that convey a distinction between, and understanding of both, enduring psychological traits and temporary states are scored high. The Affect Tone of Relationship Paradigms scale assesses the affective coloring of expectations and attributions an individual makes about his or her world. This can range from malevolent representations involving gratuitous violence to positive representations with benign and enriching interactions. The Capacity for Emotional Investment in Relationships scale ranges from viewing others as instruments for gratification, security, and comfort on the lowest level to a capacity to form deep, committed relationships in which others are valued for their unique qualities on the highest level. The Investment in Moral Standards scale measures one's moral development. At the lowest level moral development is characterized by a disregard for social norms or (slightly more developed) a ridged adherence to such norms without understanding their importance. At the highest level, moral standards are characterized by both an understanding of moral standards and a commitment to others' well-being. Finally, the Understanding of Social Causality scale assesses the complexity and maturity of one's inferences as to why people behave as they do. At the lowest level these inferences are characterized by idiosyncratic, illogical, and unlikely attributions. At the highest level, attributions involve an understanding of complex psychological motivations, internal causality (versus simple response to overt behavior), and unconscious processes.

The SCORS is typically applied to early memories or Thematic Apperception Test (TAT) stories, though it can, theoretically, be applied to any narrative data. Defenses are accounted for in various ways when scoring the SCORS. A lack of complexity, minimal description of affect, and unrealistic happy endings in TAT stories

could all reflect defensive operations and are taken into account with the SCORS (Westen, 1991).

Westen (1995) revised the SCORS. The earlier version had four subscales[12] and five levels per subscale. The current version has five subscales and seven levels per subscale. It should be noted that much of the validity research on the SCORS that follows (those studies published before 1995), was conducted with the older version. However, the revised version was expected to have as good or better psychometric properties (Weston, 1995).

Relationship with other measures of object relations and mental health. With regard to validity, the SCORS has been shown to correlate with other tests of ego development and object relations such as the Loevinger's (1976) Sentence Completion test (r = .35; Barends, Westen, Leigh, Silbert, & Byers, 1990) and the concept of the object (Blatt, Brennis, Schimek, and Glick, 1976) on the Rorschach (r ranges from .24 to .49 for relevant subscales; Hibbard, Hilsenroth, Hibbard, & Nash, 1995). In terms of assessing mental health, the SCORS has been shown to differentiate between psychiatric and healthy populations (Hibbard et al., 1995).

The SCORS as a measure of normal development and psychopathology. Normal development is also reflected by the SCORS. Children in the 2nd, 5th, 9th, and 12th grade were compared on the subscales of the SCORS. Results demonstrated that, as expected, representations become increasingly more complex between younger and older children. There is also an increase in the capacity for investment in relationships, moral standards, and in the understanding of social causality. Developmental differences were apparent

[12] The Capacity for Emotional Investment in Relationships and Investment in Moral Standards scales had been combined in the earlier version.

across ages on all subscales of the SCORS except Affect-Tone (Westen, Klepser, Ruffins, Silverman, Lifton, & Boekamp, 1991); for the latter, developmental changes are not necessarily expected; instead it may be most relevant to assessing negative events.

Thus, the SCORS has also been used to examine effects of negative developmental events. Maternal psychiatric illness, maternal alcohol abuse, prolonged separation from one's mother, and neglect had a pervasive, deleterious effect on all subscales of the SCORS, but especially on Affect-Tone (Westen, Ludolph, Block, Wixom, & Wiss, 1990). Also, such negative developmental events had a greater impact on functioning if they occurred earlier as opposed to later (preschool age and earlier versus school age). In sexually abused children and adolescents (age 5-16), the SCORS revealed simplistic and primitive characterizations of people, negative affect tone, a poor ability to invest in relationships and illogical attributions as compared to age-match controls with no history of abuse (Freedenfeld, Ornduff, & Kelsey, 1995; Ornduff, Freedenfeild, Kelsey, & Critelli, 1994).

Finally, The SCORS has demonstrated distinctions in object relations between psychological disorders. Westen, Lohr, Silk, Gold, and Kerber (1990) used the SCORS to compare patients diagnosed with borderline personality disorder, depression, and controls. Borderline patients scored lower on all subscales than normals and lower than the depressed group on Affect Tone (suggesting more malevolent relationships). Westen, Ludolph, Block, Wixom, & Wiss (1990) replicated these findings with an adolescent group of borderline patients, depressed patients, and controls who were matched on age, sex, and race. They obtained similar results for Affect Tone but results for the borderline

group were mixed on the Capacity for Emotional Investment in Relationships and Understanding of Social Causality scales.

The scale's construct validity has been tested in several areas including differential diagnosis (differentiates borderline, narcissistic, depressed, and normal populations with both adults and adolescents [Westen, et al., 1990; Ludolph et al., 1990; Gutin, 1997; Malik, 1992; Ackerman, Clemence, Weatherill, & Hilsenroth, 1999]). Although no published studies have used the SCORS specifically as a measure of mental health that circumvents defenses, its ability to tap several important areas of object relations in a projective manner makes it one of very few tests currently available that appears suitable toward this end, for the present study.

Despite support for the usefulness of the SCORS in terms of research on both normal development and psychopathology, the validity data suffer from a handful of shortcomings. The studies on normal development were uniformly cross-sectional. Such research with the SCORS would benefit from a longitudinal design. Also, several of the studies on negative developmental events used retrospective, self-reported history, which could easily be influenced by one's mental representations of those events. Nevertheless, The SCORS' ability to assess developmental capacities with projective material in an observer-rated measure (and in a timely fashion) made it the most reasonable alternative to an objective measure of mental health for the present study.

With regard to reliability, inter-rater agreement (assessed by interclass correlations) for the SCORS scales have been measured at .90 (Westen, Ludolph, Lerner, Ruffins, & Wiss, 1989). Internal consistency of the scales is rarely reported, but two studies found internal consistencies to be between .56 and .77 (Barends, Westen, Byers,

Leigh, & Silber, 1990; Westen, Lohr, Gold, & Kerber, 1990). The lowest scores (<.60) were found with the Affect Tone subscale.

Verifying reliability of the SCORS in the present study. Although the SCORS has been considered well validated compared to other projective measures of mental health, some reliability data on the measure remain questionable (Lilienfeld et al., 2000). Therefore, reliability data for this sample are detailed here. Interrater reliability was calculated using intraclass correlations (ICCs). The ICC weights agreed upon scores that are farther from the mean, and is thereby less inflated by chance agreement than r^2 (Shrout & Fleiss, 1979). An ICC above .80 is considered adequate for research measures (Shrout & Fleiss, 1979).

Two coders coded SCORS TAT data for the present study. Their ability to code data validly and reliably was checked in the following ways. First, coders were instructed on the use of the SCORS by an individual trained by Westen to conduct such instruction (Robin P. Weatherill). Then, to assess coder validity, the two coders scored 30 TAT protocols included in the SCORS manual (Westen et al., 1990), which had 'correct' scores in the appendix of the manual (i.e., The scores Westen assigned to each story). Coders were blind to Westen's scores when coding practice data. ICCs were then computed between each of the two coders' ratings and Westen's ratings for each scale of the SCORS. ICCs for the four scales used in the present study are visible in Table 2. All ICCs were within the acceptable range. Second, raters double coded 20% of the protocols from this study's data set. An ICC was computed to assess their inter-rater reliability. ICCs for the four subscales that were used in this study are reported in Table 2. All ICCs were within the acceptable range.

Internal consistency of the SCORS reflects the degree to which the scores assigned to stories of a participant (within a given scale) correlate with each other. Theoretically, scores within a subscale should be highly correlated because they are intended to measure the same construct. Internal consistency is considered 'satisfactory' at alpha = .70 or higher (Nunnally, 1978). The average internal consistency of each scale is listed in Table 2.

Internal consistency of the TAT cards when used to compute SCORS totals was disappointing overall. The only scales that reached even a modest level of consistency across the TAT cards were the Complexity and Social Causality scales. The poor internal consistency in the present study as compared to past studies (Barends et al, 1990; Westen et al., 1990) may, in part be due to the fact that that past studies that reported internal consistency data have included both healthy individuals and clinical groups with relatively severe diagnoses (major depression, bipolar disorder, borderline personality). Such samples likely yielded data with a greater range of scores between individuals. The effects of 'card pull' may also have played a role in the particularly poor internal consistency of the Affect Tone and Relationships scales, in that that some cards may have pulled for a negative affective valence across individuals. However, eliminating any one TAT card from analyses did not improve internal consistency. The low internal consistency of two out of the four scales (as well as preliminary analyses that are detailed below) argued for analyzing Affect Tone and Relationships scales separately from Complexity and Social Causality scales in further analyses.

Table 2

Rater validity, inter-rater reliability, and internal consistency of the SCORS subscales.

Scale	Validity: Interrater reliability (ICC) between each scorer and Westen on practice data		Reliability: Interrater reliability (ICC) between scorers on study data	Internal consistency (α)
	Coder 1	Coder 2		
Complexity	.89	.93	.82	.67
Affect Tone	.85	.88	.83	.36
Relationships*	.83	.81	.86	.38
Social Causality	.84	.88	.85	.57

* Practice data were coded based on a slightly older scale that combined Moral Standards and Relationships scales.

Note that the SCORS Investment in Moral Standards scale is not included in Table 1. This is because 87% of stories given had no moral theme to score, and 79% of individuals had 2 or fewer stories (out of 6 total) that had a moral theme. 36% of the sample gave no stories at all with a moral theme. For this reason it appeared inappropriate to include the scale in results.

Some researchers have critically noted a lack of test-retest data available on the SCORS (Lillienfield, 2000). The present study is the only study known to the present author that reports such test-retest data on the scale. Test-retest reliability over 15 weeks for the 4-items of the SCORS used in this study combined was .22 for the treatment group and .51 for the control group. The test-retest reliability of the experimental group is likely confounded by treatment effects, but the reliability of the control group was somewhat low for a scale intended to measure characteristics that are theoretically considered to be enduring traits.

The intervention: PSY 325, Affect and Self-Esteem.

As stated earlier, a goal of the present study was to determine whether self-report mental health measures would not accurately reflect improvement after a

psychotherapeutic intervention, due to the distorting effects of defenses on such measures. In brief, for those who are excessively defensive, effective treatment should decrease defensiveness. As excessively defensive individuals becomes less defensive they should also become more aware of painful or anxiety provoking thoughts and feelings, which could translate into endorsing more psychopathology *after* an effective intervention than *before* the intervention.

Affect and Self-Esteem, an experiential course offered at Michigan State University, is considered an analog to psychotherapy, is short-term, and directly aims to make participants aware of their affects and defenses (Kaufman, 1996). The following paragraphs summarize specific aspects of this intervention that apply to the present study and what made the intervention appear to be useful for testing the hypothesis that self-report measures may not accurately reflect psychotherapeutic change due to the distorting effect of psychological defenses.

In its typical format, Affect and Self-Esteem course meets twice a week for an hour and 40 minutes, for 15 weeks. It uses a psycho-educational curriculum that involves several components including a didactic portion, small group meetings (with the same group for each discussion to enhance rapport and comfort between group members) and reaction paper writing outside of the meetings. Thus, the course is of limited duration, but relatively intensive.

The course is based on a combination of object-relations, interpersonal, and affect theories. The didactic portion and group discussions center around the use of specific tools that are presented and worked with both inside and outside of the classroom. A brief outline of the course, theory, and tools is included in appendix B. For a more in

depth discussion of the theory and tools see Kaufman (1996). One of the fundamental theoretical underpinnings of the course is rooted in powerlessness-affect-stress cycles. This concept centers on the idea that our inability to predict or control life events often leads to a feeling powerlessness, which yields negative affects (such as sadness, anger, and anxiety). In turn, the attempt to escape this affect (with any of the various methods of defense ranging from repression to suppression) results in what Kaufman terms "backed-up affect," which creates more generalized stress (Kaufman, 1996, p. 248.). Students are required to write about a powerlessness-affect-stress cycle in their own life to facilitate their ability to recognize and bring to consciousness their own painful affects and means of coping with such affects. Kaufman explains that writing about aspects of this cycle allows participants to "consciously observe the sequence of coassembled inner states: activator – affect – consequence" (1996, p. 249). Another important aspect of Kaufman's program involves participants' learning, using, and writing about specific tools used to cope with difficult situations and painful or anxiety provoking feelings (1996). Such tools involve using what others would term highly adaptive defenses, including the use of humor, sublimation, and anticipation.

In short, several dimensions of Kaufman's intervention are designed to bring affects into awareness and to curtail extreme defenses against such painful or anxiety-provoking feelings, in a time-limited treatment, making it a useful test for the effects of defenses on treatment outcome measurement.

Procedure

On the first and last day of each course, participants were handed packets including a consent form, lined paper for TAT responses, the State-Trait Anxiety

Inventory, the Beck Depression Inventory, the Marlowe-Crowne Social Desirability Scale, and the MMPI Denial scale mixed in with 30 other randomly selected MMPI items. They also were given a packet of 6 TAT card copies (the minimum number suggested by Westen [1991]) enclosed in an envelope. Measures required a total of approximately 50 minutes to 1 hour to complete. Participants were asked to read the consent form and sign it if they choose to participate. Participants were verbally reminded that they were free to discontinue participation at any time, without penalty. Those who participated were entered in a raffle to win $100.00.

After consent was obtained, Participants were given standardized TAT instructions as per Murray (1968) and instructed to open their packet of cards and look at the first page. Participants were given 3 minutes to respond to each card. They were also given the following queries at the specified time intervals, for each card: "What happened before the scene on the card (at 45 seconds)? What are the people in the card thinking and feeling (at 90 seconds)? What happened after the scene on the card – how does the story end (at 135 seconds)?" Once TAT responses were completed, participants were asked to complete questionnaire measures. On the last day of each course, participants were asked to complete all measures again in the same manner as above.

To control for the effects of "card pull" (the possibility that some TAT cards would draw for lower scoring themes than others), card presentation was randomized as follows. Each participant responded to 12 cards in total – 6 at pretest and 6 at posttest. Labeling the cards 1 through 12, half of each group responded to cards 1-6 at pretest and 7-12 at posttest. The other half responded to cards 7-12 at pretest and 1-6 at posttest.

Data Reduction

Three factors were of interest in the following analyses: defensiveness, self-report mental health, and observer-rated mental health. Several preliminary analyses were performed to determine whether some measures might be more robust than others in quantifying the factors in question and, therefore, to determine which combinations of scales would be most appropriate for hypothesis testing. Where composite variables were created out of separate scales, alphas are reported at the conclusion of this section.

Relationship among defensiveness measures: The MMPI Dn scale and the MCSD scale correlated significantly (r=.37, p = .001). Though significant, this correlation was relatively weak given that the measures are intended to assess the same factor (defensiveness) in this study. Considering these measures as items in a scale, they achieved an alpha of .52, which was poor. Therefore it appeared more appropriate to quantify The MCSD scale and the MMPI Dn scale separately rather than combine them in further analyses.

Relationship among self-report mental health measures: Correlations among self-report mental health measures were between r=.55 (p<.000) and r=.68 (p<.000). When considering the BDI, STAI state and STAI trait as items in a scale, alpha reached .81, which was adequate. Eliminating any one scale decreased alpha. Therefore it was decided to combine these measures in further analyses.

Table 3

Correlations among self-report mental health measures

	BDI	STAI state
STAI state	.55***	
STAI trait	.68***	.58***

*** indicates a significant finding at p < .000

Relationship among observer-rated mental health variables. Correlations among the SCORS measures varied. Complexity and Social Causality were acceptably correlated (r=.58, p<.000) as were Affect Tone and Relationships (r=.47, p<.000). However Affect Tone correlated poorly with the Complexity (r=.12, p=.32) and Social Causality (r=.19, p=.10), and Relationships correlated only modestly with Complexity (r=.26, p=.03) and Social Causality (r=.35, p=.01). A reliability analysis was then conducted on the 4 SCORS subscales. Alpha of the four scales (treated as 4 items) was .66. However, when subscales were split into the two groups (Complexity + Social Causality and Affect Tone + Relationships, alphas were .73 and .68 respectively). The reason for this discrepancy may be that Affect Tone is the only scale that is unrelated to development. That is, the affective valence of individuals' attributions is not expected to develop from malevolence to benevolence in the same way as understandings of individual's internal states and attributions for people's actions are expected to become more sophisticated over the course of development. However, the Relationships scale may also be highly affected by the affective valence of stories. It was exceedingly rare that a story involved highly negative affect tone, but also included a description of deep, committed, and interdependent relationships. On the other hand, several stories involved highly negative affect tone, but still made sophisticated attributions about people's internal states and reasons for their actions. Therefore, it was decided that Complexity and Social Causality would be analyzed as a composite score separately from Affect Tone and Relationships (also analyzed as a composite score) based on both empirical and theoretical grounds. Where composite subscales are analyzed as such, they are termed "Complexity + Social Causality" and "Affect-Tone + Relationships."

Relationship between defensiveness measures and self-report mental health. For the next preliminary analysis, correlations were calculated on defensiveness measures and the self-report mental health composite measure. The correlations generated are visible in Table 4. As expected, defensiveness measures generally had a significant, positive relationship with self-report mental health. Thus, greater defensiveness was associated with greater self-report of mental health. The MMPI Dn scale appeared to have a stronger relationship with the composite self-report mental health variable. It is unclear why this would have been the case, but this finding, combined with the low correlation between defensiveness measures, supported the quantification of defensiveness measures separately rather than combined.

Table 4

Relationship between defensiveness and self-report mental health

	self-report mental health
Dn	.47***
MCSD	.18*

* indicates $p < .05$
*** indicates $p < .001$

Relationship between observer-rated mental health and defensiveness. Correlations were conducted on observer-rated mental health and defensiveness measures. The matrix generated is visible in Table 7. Results were generally as predicted for the Complexity + Social Causality composite of the SCORS. That is, higher defensiveness was correlated with poorer observer-rated mental health. However, this was not found in relation to the Affect Tone + Relationships composite. This could have occurred because high scores on this composite are associated with a positive affective valence, and such a valence

could be influenced by defensiveness. In support of this, the positive correlation between the MCSD scale and Affect Tone + Social Causality composite approached significance (p=.08). Thus, high defensiveness on the MCSD scale was marginally related to higher (healthier) scores on the Affect Tone + Relationships subscale of the SCORS. This finding further supported the separate quantification of SCORS composite variables.

Table 5

Relationship between observer-rated mental health and defensiveness

	SCORS Complexity + Social Causality	SCORS Affect Tone + Relationships
Dn	-.22*	.02
MCSD	-.25**	.16

* indicates $p < .05$
** indicates $p < .01$

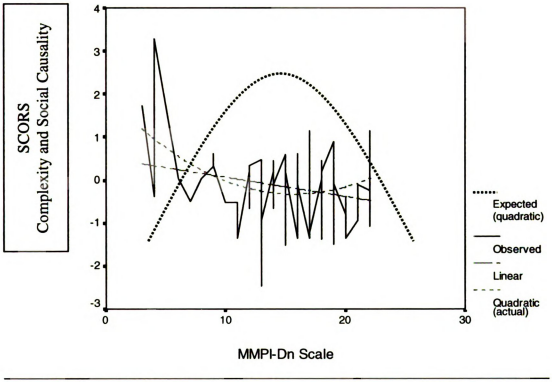
As stated above, it was posited that defensiveness measures and the SCORS could have been curvilinearly related such that very high or very low scores on the defensiveness measures would be associated with poor observer-rated mental health and midrange scores would be associated with higher observer-rated mental health. This particular curvilinear relationship would be reflected in a quadratic linear equation. To test this, the next preliminary analyses used higherarchical linear regression. The dependant variable was the Complexity + Social Causality Composite[13]. In the first regression model, The MCSD scale was entered in block 1 and the quadratic term of this variable was entered in block 2 (MCSD**2). The linear term of the MCSD scale was significant (r^2 = .06, p = .03) and the quadratic term did not add incrementally to the linear term (r^2 = .07, p = .73). In the second regression model the MMPI-Dn scale was

[13] Results for the Affect-Tone + Relationships composite are not described here because the correlation between this composite and defensiveness measures suggested a regression on this variable would not clarify the anticipated curvilinear relationship between observer-rated mental health and defensiveness.

entered in block 1 and the quadratic term of the MMPI-Dn scale (MMPI-Dn**2) was entered in block 2. The linear term was marginally significant ($r^2 = .04$, $p = .05$) and the quadratic term accounted for variance beyond the linear term ($r^2 = .12$, $p = .02$). However, the quadratic term was in the opposite direction than had been anticipated.

Figure 1

Curve fit estimation: MMPI Dn scale to SCORS Complexity + Social Causality



As is evident in Figure 1, the effect was primarily due to 1 participant who exhibited very low defensiveness and obtained an unusually high score on the Complexity + Social Causality composite variable (greater than 3 standard deviations above the mean). When this individual was eliminated from the analysis, the curvilinear term no longer significantly accounted for variance beyond the linear term. In short, the

suggestion that very high or very low scores on the defensiveness measures would be indicative of poor observer-rated mental health and midrange scores would indicate higher observer-rated mental health was not supported by any combination of the defensiveness and observer-rated mental health measures.

Relationship between SCORS and self-report mental health. Correlations were conducted on the observer-rated mental health composite variables and the self-report mental health composite. Neither observer-rated measure was significantly correlated with the self-report mental health composite (see Table 5). A weak relationship between self-report mental health and observer-rated mental health was expected, but the lack of relationship was unexpected.

Table 6

Correlations between observer-rated mental health and self-report mental health

	Complexity + Social Causality	Affect Tone+ Relationships
Self-report mental health	-.16	.02

Conclusions based on preliminary analyses. Preliminary analyses indicated that the self-report scales held together well when considered as items on a scale. A **composite self-report mental health** (61 items: alpha = .96) variable was therefore used in further analyses. This was a combination score of the State scale of the STAI, the Trait scale of the STAI, and the Beck Depression Inventory. For each participant, the three measures were converted to z-scores (based on combined group means and standard deviations) and then summed to form the composite self-report mental health score. Separate variables were created from time one and time two.

Reliability data and preliminary analyses indicated that the SCORS measures were best combined into two separate variables: **Complexity + Social Causality** (12-items: alpha = .76) and **Affect Tone + Relationships** (12 items: alpha = .58). These variables were considered separately in all analyses. Each composite scale consisted of participant's 12 scores (6 TAT cards x 2 SCORS subscales), which were summed to obtain the two observer-rated mental health scores. Separate scores were created for each variable from time one and time two.

Preliminary analyses demonstrated that the MMPI Dn scale and the MCSD scale correlated poorly and did not hold together when analyzed as items on a scale. Therefore, these scales were analyzed separately. For each participant, the total raw score on each measure was converted to a z-score to form the **MMPI-Dn scale** score and the **MCSD scale score**. Separate variables were created from time one and time two for each scale.

Sample Size and Power Analysis

Determining an adequate number of participants for sufficient power in this study was difficult since few studies of this kind have been conducted in the past and the pioneering study in this area (Shedler, Mayman, & Manis, 1993) did not publish standard deviations (making it impossible to determine effect sizes). However, determining the effect sizes of previous studies that used the SCORS, as well as effect sizes in the literature on interventions similar to the group format of the Affect and Self-Esteem course, yielded a reasonable estimation of the effect size that could be expected in this study.

Previous literature on the SCORS (Westen et al., 1990) has demonstrated between-group effects ranging from $\underline{f}$ = .55 to .68, when comparing individuals with

various psychopathologies to controls. Such effect sizes are in the "large" range according to according to Cohen (1992) and correspond to d between 1.0-1.4. The SCORS has also demonstrated "large" effect sizes (r^2 = .16) when used as a measure of psychological development among normal children and adolescents (Westen et al, 1991). Meta-analyses of the effectiveness of group psychological treatments (McDermut, Miller, & Brown, 2001; Tillitski, 1990; McRoberts, Burlingame, & Hoag, 1998) found effect sizes in the large range (d = .90 to 1.82) by Cohen's (1992) standard[14]. Thus, it is reasonable to estimate that the present study, using the SCORS to evaluate the effectiveness of an analog to group treatment, would yield medium to large effect sizes if it was to yield effects of clinical interest. With 31 participants in each cell (the minimum number used), the present study had .92 power to detect medium effect sizes (f = .25) in main effects and .86 power to detect the same effect sizes in interactions in a 2x2 ANOVA with no covariates. Smaller effect sizes than those detectable are of questionable clinical relevance in this literature.

[14] These studies included a combination of observer report and self-report measures.

RESULTS

Preliminary Description

As data are presented in the text, F values and p are listed where appropriate, as is partial eta-squared (η^2). η^2 is an estimate of effect size, describing the proportion of total variance attributable to a factor (Cohen 1992). It is interpreted in a similar manner as r^2 is interpreted.

Group differences before treatment: Table 6 demonstrates that the treatment and control group were well matched in terms of age, grade, sex, and ethnicity. Table 6 also reveals that the control group scored significantly higher than the treatment group on the SCORS Affect Tone + Relationships at time 1. These results suggest that the treatment group made less benign attributions about relationships and viewed relationships in a more need-gratifying versus interdependent way.

Table 7

Descriptive data for treatment and control groups at time 1.

	Treatment Group	Control Group	p value
Total N	31	45	--
Age (M/SD)	21.8/2.45	21.8/2.6	.99
Grade (M/SD)	3.8/.50	3.9/.25	.33
Sex (Male/Female)	6/25	10/35	.76
Ethnicity	--	--	.72
Caucasian (Frequency)	26	37	--
African-American (Frequency)	2	1	--
Hispanic/Latino (Frequency)	2	6	--
Asian (Frequency)	1	1	--
Self-Report Mental Health Composite (M/SD)	-.14/.80	.12/.82	.17
MCSD (M/SD)	14.6/5.8	15.3/5.6	.61
MMPI Dn Scale (M/SD)	14.5/4.7	15.5/3.9	.30
SCORS Complexity + Social Causality (M/SD)	-.24/1.05	-.10/.70	.50
SCORS Affect Tone+ Relationships (M/SD)	-.32/.80	.10/.74	.02*

* indicates a significant finding at $p < .05$

Covariates. Correlations were conducted between variables of interest and demographic data to determine if covariates exist in the data. The correlation matrix generated is visible in Table 3. Several demographic variables correlated significantly with variables of interest. Age correlated with the MMPI Dn scale (r = .27, p = .02), such that older age was related to higher defensiveness. Visual inspection of the data revealed that the significant relationship was largely due to 3 outliers -- no significant relationship was present when the 3 subjects above age 30 were removed from the analysis (r = .19, p = .10). Nevertheless, age was covaried in further analyses of the MMPI Dn scale.

Table 8

Correlations between variables of interest and demographic data (r)

	Age	Grade	Sex	Ethnicity
Self-report mental health	.09	-.13	.08	-.15
MCSD	.14	.02	-.11	.19
MMPI Dn scale	.27*	.05	.02	.07
SCORS Complexity + Social Causality	-.21	.07	.23*	-.28*
SCORS Affect Tone + Relationships	-.10	-.00	-.00	-.29*

* indicates a significant finding at $p < .05$
** indicates a significant finding at $p < .01$

Although there were no significant correlations between age and SCORS composites, the direction of the relationships between them was in opposite direction of what would be anticipated for all SCORS variables (i.e. higher age was correlated with poorer observer-rated mental health rather than better observer-rated mental health as would have been expected). Eliminating outliers (based on age and defensiveness) did not change this relationship. Based on scatter plots, it appears that younger age correlates with greater standard deviation of the SCORS composites as well.

Women scored significantly higher than men on SCORS Complexity + Social Causality (r=.23, p=.04). This may indicate that the female participants' representations

of people were better distinguished, more multidimensional, and better integrated than those of male participants. It is also possible that the women had higher verbal facility than men, which likely would have yielded more complex representations of people and the reasons for their actions. Sex was covaried in subsequent analyses.

Finally, ethnicity correlated with both the SCORS Complexity + Social Causality composite and the SCORS Affect-Tone + Relationships composite. Those who were not Caucasian scored lower on both scales. Ethnicity was covaried with further analyses of the SCORS variables.

Order effects. Participants where administered one of two sets of 6 TAT cards pretreatment and the other set of 6 cards post treatment (set of cards A and set of cards B). The order of administration (AB or BA) was randomized across groups as is made visible in Table 9.

Table 9

Order of TAT card administration

	Treatment group	Control group
Order AB	15	24
Order BA	16	21

The order in which TAT cards were administered correlated with scores on the Complexity + Social Causality composite of the SCORS (r =.31, p=.007) at time one, but not at time two. Apparently cards in the B group pulled for poorer scores than cards in the A group, but only at time one. It is unclear why this would be the case. Order effects were covaried in further analyses involving the SCORS Complexity + Social Causality scales.

Hypothesis 1.

An interaction was hypothesized in which self-report mental health would correlate with observer-rated mental health at low levels of defensiveness, but not at higher levels of defensiveness. Hierarchical linear regression was used to determine the degree to which defensiveness moderated this relationship between self-report mental health and observer rated mental health.

Analysis 1a: The analysis was first conducted with Complexity + Social Causality as the dependant variable. Composite self-report mental health was entered in block one, the MMPI Dn scale in block 2, and self-report mental health*MMPI Dn in block 3. As shown in Table 9, Block 3 was nonsignificant and accounted for virtually no variance beyond block 2. Thus, the predicted interaction was not observed in relation to the MMPI denial scale and Complexity + Social Causality. When the same regression was run with the MCSD scale instead of the Dn scale, block 2 accounted for a significant proportion of the variance over model 1, but block 3, again, accounted for no variance over block 2 (summarized in Tables 10-13). The same regressions were then run with the other composite of the SCORS (Affect Tone + Relationships) as the dependant variable. Again, block 3 was nonsignificant (r^2=.-04, p=.81) and accounted for little variance over block 2 (r^2=.-03, p =.89). Finally, the regression was run with the MCSD scale as the moderator. Block 3 was nonsignificant (r^2= .-05, p =.20) and accounted for little variance beyond block 2 (r^2= .-03, p =.16).

Although the MMPI-Dn scale better predicted Complexity + Social Causality than self-report mental health, neither defensiveness measure moderated the relationship between self-report and observer-rated mental health. Thus, hypothesis 1 was not

supported. Analysis 1b was moot because preliminary analyses determined that neither defensiveness measure had the predicted quadratic relationship with observer-rated mental health. Therefore, it could not be expected that running the same analysis with the quadratic equation of the defensiveness measures would better predict observer-rated mental health measures.

Table 10

Hypothesis 1: I.V. = MMPI Denial, D.V. = Complexity + Social Causality.

	Beta	r^2	r^2 change	p-value
Block 1: SRMH	-.16	.03	.027	.16
Block 2: SRMH	-.08			.56
MMPI-Dn scale	-.18	.05	.026	.16
Block 3: SRMH	-.40			.38
MMPI-Dn scale	-.16			.23
SRMH * MMPI Dn scale	.32	.06	.007	.45

Table 11

Hypothesis 1: I.V. = MCSD, D.V. = Complexity + Social Causality.

	Beta	r^2	r^2 change	p-value
Block 1: SRMH	-.16	.03	.027	.16
Block 2: SRMH	-.12			.30
MCSD scale	-.23	.08	.051	.05*
Block 3: SRMH	-.15			.62
MCSD scale	-.23			.05*
SRMH * MCSD scale	.03	.08	.000	.91

Table 12

Hypothesis 1: I.V. = MMPI Denial, D.V. = Affect Tone + Relationships

	Beta	r^2	r^2 change	p-value
Block 1: SRMH	.02	.00	.000	.90
Block 2: SRMH	.01			.97
MMPI-Dn scale	.02	.00	.000	.87
Block 3: SRMH	-.11			.80
MMPI-Dn scale	.03			.83
SRMH * MMPI Dn scale	.12	.00	.001	.79

Table 13

Hypothesis 1: I.V. = MCSD, D.V. = Affect Tone + Relationships

	Beta	r^2	r^2 change	p-value
Block 1: SRMH	.02	.00	.000	.90
Block 2: SRMH	-.02			.90
MCSD scale	.16	.02	.024	.18
Block 3: SRMH	-.41			.20
MCSD scale	.14			.26
SRMH * MCSD scale	.43	.05	.024	.18

Hypothesis 2a

It was hypothesized that after a psychotherapeutic intervention, individuals in the treatment group would demonstrate improved observer-rated mental health but not self-report mental health as compared to controls. Thus, a 2*2*2 ANOVA (group [treatment vs. control] by mental health report type [self-report vs. Complexity + Social Causality of the SCORS] by time [pre vs. post]) was run to determine if the above relationship would manifest in a three-way interaction (group * mental health report type* time interaction). As predicted, the 3-way interaction was significant F (1, 50) =6.09, p=.02, η^2=.11. However, 2*2 ANOVAs (self-report mental health * group and observer-rated mental health * group) indicated that the direction of relationships was not as predicted (see Figure 2 and Figure 3).

In terms of Complexity + Social Causality, an overall increase (improvement) in scores was exhibited between time 1 and time 2, F (1, 50) = 10.22, p = .002, η^2 = .17, with a non-significant interaction between groups. However, with regard to self report mental health, a significant 2-way interaction was found such that the treatment group improved between time one and time two whereas the control group declined, F (1,50) = 6.80, p = .01, η^2 = .12. Paired-sample T-tests were then conducted to determine which

changes were significant. Changes in self-report mental health and the Complexity + Social Causality subscales of the SCORS reached significance for the control group, t=(26) 2.33, p =.03 and t= (26) -3.25, p =.003 respectively) but not for the treatment group.

Figure 2

Hypothesis 2a: actual group * mental health type * time interaction.

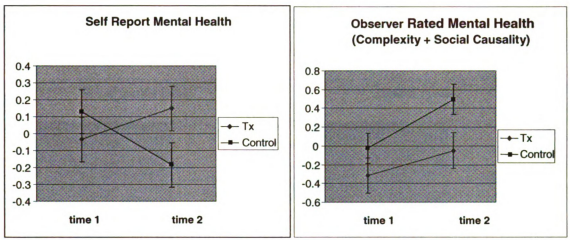
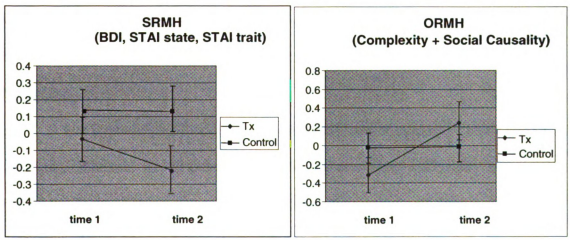


Figure 3

Hypothesis 2a expected interaction.

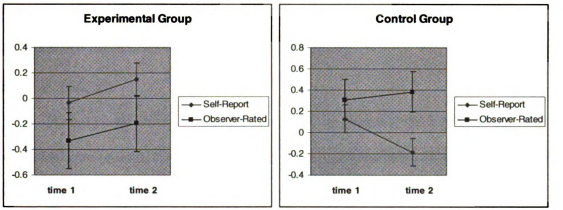


The same 2*2*2 ANOVA was then computed with the other composite SCORS subscale

(Affect Tone + Relationships). Only the mental health * group interaction was significant, $\underline{F}$ (1,50) = 6.42, $\underline{p}$ = .01, η^2 = .11). As is visible in Figure 3, 2*2 ANOVAs indicated that the experimental group displayed non-significant improvement on self-report mental health and observer-rated mental health whereas the experimental group displayed a marginally significant interaction between observer-rated and self-report mental health measures, $\underline{F}$ (1, 26) = 3.90, $\underline{p}$ = .06, η^2 =.13.

Figure 4

Hypothesis 2a: group * mental health type* time interaction.



T-tests demonstrated a non-significant improvement in terms of observer-rated mental health, but significant decline in terms of self-report mental health, $\underline{t}$ (26) = 2.33, $\underline{p}$ = .03. Hypothesis 2a was therefore not supported.

Hypothesis 2b.

It was hypothesized that defensiveness would exert a suppressor effect on self-report mental health. Recall that defensiveness was associated with higher self report mental health (see Table 6 above). In light of this, if defensiveness decreased between time one and time two (as had been expected in the treatment group), then covarying defensiveness could make visible an improvement in self-report mental health at time 2

over time 1 in the treatment group. As noted in Hypothesis 2b above, a paired samples t-test failed to demonstrate an increase in self-report mental health between time 1 and time 2 for the treatment group. Of course, the same was observed with a repeated measures ANOVA, $\underline{F}$ (1, 24) = 1.85, $\underline{p}$ = .19, η^2 = .07. However, the repeated measures ANOVA demonstrated a significant increase in self-report mental health between time 1 and time 2 when the MMPI Dn scale at time 1 was covaried, F (1, 24) = 6.73, $\underline{p}$ = .02, η^2 = .22, which was consistent with the hypothesis. This did not emerge when the MCSD scale was covaried, $\underline{F}$ (1, 24) = 1.09, $\underline{p}$=.31, η^2 = .05). Thus, Hypothesis 2b was partially supported. The predicted relationship was found when the MMPI Dn scale was covaried, but not when the MCSD scale was covaried.

Post hoc analyses

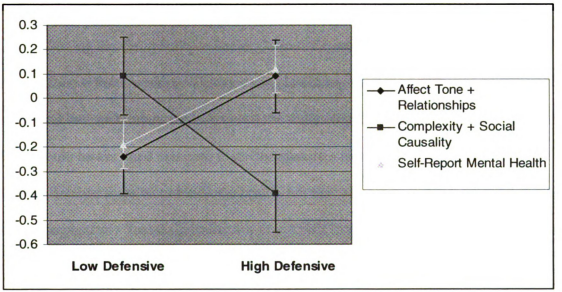
Although preliminary analyses suggested the expected relationships between variables were present, specific predictions related to hypothesis 1 were not supported. Several post hoc analyses were conducted to determine what may have accounted for the unanticipated relationships between observer-rated mental health and other variables.

Defensiveness groups. Splitting participants into groups based on defensiveness and comparing them on self-report and the SCORS measures (as had been planned if hypotheses were supported) remained likely to clarify relationships between these factors. In a 2*2 repeated measures ANOVA, a significant group (low defensive/high defensive) by mental health (self-report versus Complexity + Social Causality) interaction was found, both when the median split was conducted on the MCSD scale , $\underline{F}$ (1, 50) = 10.4, $\underline{p}$=.002, η^2=.12 and when the median split was conducted on the MMPI-Dn scale $\underline{F}$ (1,50) = 4.48, $\underline{p}$ =.03, η^2 = .07). As had been predicted in hypothesis 1, high

defensiveness was associated with high self-report mental health, but low observer rated mental health. However, low defensiveness predicted the opposite relationship – low self-report mental health and high observer-rated mental health. This was true for both the MMPI-Dn scale and the MCSD scale. This relationship had not been expected and accounts for the negative findings of the linear regression run in hypothesis 1, because self-report and observer rated mental health were negatively related at both high and low levels of defensiveness (thus, defensiveness did not moderate the relationship between them).

Figure 5

Self-report and observer-rated mental health at high and low MCSD scores



Given the above results, it appeared unlikely that Complexity + Social Causality or Affect-Tone + Relationships could be considered accurate measures of mental health when considered in isolation. Affect Tone + Relationships did not appear to account for defensiveness. Complexity + Social Causality, although related to defensiveness in the expected manner, was not related to low levels of self-report mental health in the

expected manner (i.e., it was not predicted that individuals would be judged as healthy based on any SCORS measure, but subjectively endorse psychopathology, as is evident above).

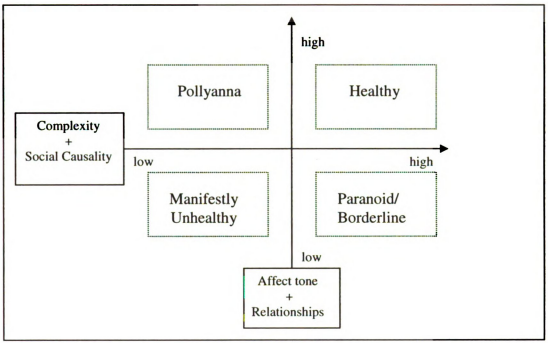
Typological analysis of SCORS composites. It appeared possible that relationships between SCORS variables could be more complex than initially proposed. Rather than higher scores indicating health across subscales, variations between scales may have had important implications for psychological functioning. Specifically, combinations of high and low scores on the SCORS could be as unhealthy as or less healthy than low scores *across* the SCORS subscales. See appendix G. for an expanded description of the rational behind the typological analysis of the SCORS performed in the following post hoc analyses.

Setting the Complexity + Social Causality composite average perpendicular to the Affect Tone + Relationships subscales and classifying the 4 corresponding groups suggested that such interactions between the scales could, theoretically, exist (See Figure 5). It could be assumed that both high Complexity + Social Causality and high Affect Tone + Relationships would indicate health, but that some individuals who fall into the high Complexity + Social Causality but low Affect Tone + Relationships quadrant could exhibit borderline functioning (see appendix G for expalination). Similarly, many individuals falling in the opposite quadrant (low Complexity + Social Causality combined with high Affect Tone + Relationships) might be highly defensive. There would be reason to believe that some individuals in the "defensive" and "borderline" quadrants may be as "psychologically unhealthy" or less healthy than manifestly unhealthy individuals who score lower on the SCORS when scales are simply combined

and added together. If this were the case, one might expect individuals in the "defensive" quadrant to obtain the highest scores on defensiveness measures, individuals in the "borderline" quadrant to obtain score the lowest, and the other two groups to score in between.

Figure 6

SCORS classifications based on quadrants.

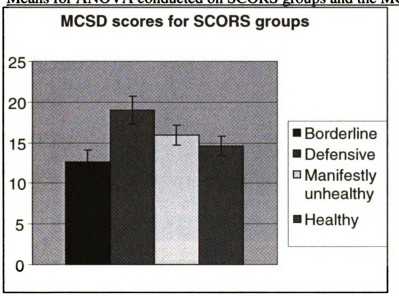


After individuals scoring within ½ standard deviation of the mean of both scales were eliminated from analyses, one-way ANOVAs were conducted on the four SCORS categories based on the above quadrants (hence forth referred to as "SCORS groups") and defensiveness scores (MCSD scales and MMPI-Dn scale). A main effect for group was found (F=2.84, p=.04, η=.12) and planned comparisons demonstrated a significant difference between the borderline and defensive SCORS groups (F=6.13, p=.007) on the MCSD scale scores. As expected, the group classified as defensive (high Affect Tone +

Relationships and low Complexity + Social Causality) were, in fact, the most defensive group, as measured by the MCSD scale. The group classified as borderline (low Affect Tone + Relationships and high Complexity + Social Causality) scored lowest (least defensive) on the MCSD scale. The manifestly unhealthy group (low Affect Tone + Relationships and low Complexity + Social Causality) and healthy group (high Affect Tone + Relationships and high Complexity + Social Causality) scored between the defensive and borderline groups, but did not differ significantly from either.

Figure 7

Means for ANOVA conducted on SCORS groups and the MCSD scale

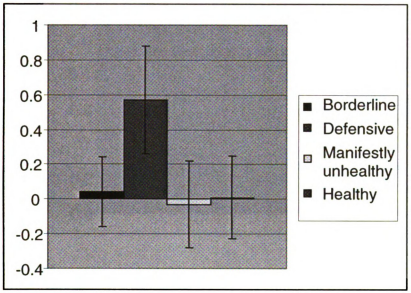


When the same ANOVA was conducted with the MCSD scale replaced by the MMPI-Dn scale, results were non-significant (F= .48, p=.70) but means for the 4 groups were in the same direction – the defensive group scored highest on the MMPI-Dn scale, the borderline group scored the lowest, and the depressed and healthy groups scored between the defensive and borderline groups. When comparing SCORS groups in terms of self-report mental health measures, results were not consistent with expectations. Diagnostic grouping suggested that the groups classified as borderline and depressed

would score the lowest on self-report measures of mental health and that the groups classified as healthy and defensive would score the highest. However, results from a one-way ANOVA demonstrated a non-significant difference between groups (F=2.0, p=.12, η^2 = .08). Most of this marginal effect was due to defensive individuals scoring highest on self-report mental health. Means for the groups classified as borderline, depressed, and healthy were nearly identical.

Figure 8

Means for SCORS groups based on self-report mental health



The same analysis conducted on the BDI alone, demonstrated means in the predicted direction, but results remained non-significant (F = 1.64, p=.18, η^2= .07). However, when defensive and healthy groups and borderline and depressed groups were combined to achieve greater power, results achieved marginal significance (F = 3.83, p=.05, η^2= .06).

Given the above, albeit modest, support for validity of groupings based on interactions of SCORS subscales, it appeared as though it might be useful to determine if

different SCORS groups predicted different changes in defensiveness and self-report mental health over the course of the treatment. However, group sizes appeared too small to make valid comparisons between SCORS groups within treatment and control groups (see Table 14).

Figure 9

Means For SCORS Groups based on the BDI.

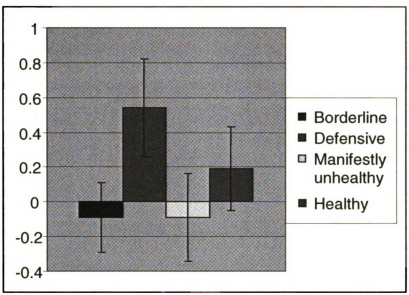


Table 14

Number of participants in SCORS groups for pre-post analyses.

	Treatment group	Control group
Healthy	3	10
Depressed	9	3
Defensive	3	4
Borderline	5	5

As could be expected, differences between SCORS groups within treatment and control groups over time were non-significant in terms of self-report mental health (F=.61, p=.61, η^2= .05) and BDI scores (F=.95, p=.44, η^2= .07). Although overall differences were found between SCORS groups in terms of the MCSD (F=3.61, p=.02,

η^2= .24), a three-way interaction that would have indicated differences between SCORS groups within the treatment or control group over time was also non-significant (F=.41, p=.75, η^2=.04). Main effects of the SCORS groups on the MMPI-Dn scale and a three-way interaction between SCORS groups within treatment and control groups over time were also non-significant (F=1.16, p=.34, η^2=.09 and F=.25, p=.86, η^2=.02, respectively).

DISCUSSION

The purpose of this study was to examine whether the validity of self-report measures of mental health would be brought into question by the effects of psychological defenses on responses to such measures in a treatment study. Prior studies have demonstrated psychological and physiological costs for excessive defensiveness such as poor global functioning, psychological illness and personality pathology (Perry & Cooper, 1989; Vaillant & Drake, 1985). Excessive defenses have also been linked to health problems, heath center visits, and immune functioning (Pennebaker, 1985; Pennebaker, 1989; Pennebaker & Beall, 1986; Pennebaker & Heeron, 1984; Pennebaker et al., 1988; Pennebaker & Susman, 1988); as well as unhealthy blood pressure spikes (Shedler et al., 1993) and rate of cancer spread (Jensen, 1987). Previous studies have also suggested that excessive defensiveness leads to a minimization of people's self-report of psychopathology (Davis, 1987; Mikulincer & Orbach, 1995; Schedler et al., 1993; Weinberg et al., 1979). Although some have suggested that excessive defensiveness could attenuate treatment-outcome results if individuals are defensive during pretreatment measurement (Shedler et al., 1993), this is the first study to test this possibility. The present study found partial support for the hypothesis that defensiveness attenuates treatment outcome measurement with self-report measures of mental health. Covarying defensiveness made apparent an improvement in self-report mental health over the course of a treatment where such improvement had not been evident without taking account of defensiveness. However, defensiveness did not moderate the relationship between self-report and observer-rated mental health measures. Several relevant aspects of findings and possible interpretations of the data are considered below.

Pretreatment Differences.

Differences between groups: Prior to the intervention, the control group scored higher (healthier) than the treatment group on the Affect Tone + Relationships scale of the SCORS. It is possible that the treatment group's tendency to make more malevolent attributions about others and to view relationships in a more need gratifying way influenced their decision to enroll in the course "Affect and Self-Esteem," which is presented to students as an experiential course designed to give them psychological tools. On all other variables the treatment and control groups appeared well matched prior to the intervention.

Demographic differences: The SCORS Complexity + Social Causality subscale correlated with sex and ethnicity, such that males scored lower than females and those of minority ethnicity scored lower than Caucasians on the measure. It is likely that women had greater verbal facility than men and that Caucasians had greater verbal facility than minorities, especially because 4 out of the 12 individuals of minority ethnicity were non-native English speakers. However, lower verbal facility did not account for minorities' lower scores on the Affect-Tone and Relationships subscales of the SCORS. It is evident, however, that those of minority ethnicity generally may have made more malevolent attributions about relationships -- perhaps due to experiences of prejudice having influenced how they perceive such relationships (see Tinsley-Jones, 2003 for a review).

Relationships between defensiveness, self-report mental health, and the SCORS.

Overall, expected relationships were present in the data in preliminary analyses. As expected, high defensiveness (on both scales) correlated with positive self-reports of

mental health and low Complexity + Social Causality subscale of the SCORS, suggesting that the defensiveness measures had some validity. Nevertheless, the expected relationships did not bear out in specific hypothesis tests.

Hypothesis 1

Despite extensive efforts to find such an effect, defensiveness simply did not moderate the relationship between self-reported and observer-rated mental health. After splitting participants into groups based on high and low defensiveness, it appeared that the high defensiveness group reported high mental health and were judged as relatively healthy based on the Affect Tone + Relationships composite, but were judged as relatively unhealthy based on the Complexity + Social Causality composite. However, low defensiveness predicted the opposite – low self-report mental health, relatively unhealthy Affect Tone + Relationships, and relatively healthy Complexity + Social Causality. This was true for both the MMPI-Dn scale and the MCSD scale. This unexpected pattern accounts for the lack of a moderator effect found for hypothesis 1 because self-report health was negatively related to Complexity + Social Causality and positively related to Affect Tone + Relationships *regardless* of the level of defensiveness.

Assuming that all measures were valid and reliable, the lack of a moderator effect of defensiveness on the relationship between self-report and observer-rated measures may indicate that self-report is just as valid as observer-report in measuring mental health in treatment outcome research. The lack of agreement between self-report and observer-rated measures could simply point to the theoretical existence of multiple and distinct facets of mental health.

It is also possible that Complexity + Social Causality or Affect-Tone + Relationships cannot be considered accurate measures of mental health in isolation, at least in a relatively healthy sample. Affect Tone + Relationships did not appear to account for defensiveness. Complexity + Social Causality, although related to defensives in the expected manner, was not related to low levels of self-report mental health in the expected manner. These findings using the SCORS as a measure of ego strength and as a counter measure to self-report mental health, were not consistent with past research that has used physiological data to test the accuracy of self-report in relation to defensiveness. Such studies have found physiological data consistent with self-report at low levels of defensiveness, but indicative of anxiety and inconsistent with self-report measures at higher levels of defensiveness (Dozier & Kobak, 1992; Hughes et al., 1994; Jorgenson, 1992; Lazarus & Alfert, 1964; Shedler et al., 1993; Weinberger, 1979).

Hypothesis 2

This study also failed to find the expected interaction effects of a psychotherapeutic intervention on self-report and observer rated mental health. As expected, no difference was found between pre and post treatment self-report mental health in the treatment group. However, contrary to expectation, no difference was found between pre and post observer-rated mental health for this group either. Further complicating the interpretation of results, the control group displayed an increase in Complexity + Social Causality between time 1 and time 2, but a decrease in self-report mental health over the same interval -- exemplifying the effect that was expected in the treatment group. It is unclear why the control group displayed an increase in Complexity between time 1 and time 2. One possible explanation for anomalous SCORS results is

that practice effects may have had a differential effect on groups. For instance, it is possible that the control group simply got better at the task after time 1 measures were administered. This may especially be the case because students had a limited time to respond to each card. Without experience, students in both groups may have limited the density of their descriptions to make sure they finished stories within the given time limit, which likely would have lowered scores on Complexity + Social Causality. Upon posttest, the control group may have exhibited a typical practice effect whereby they became more skilled at the task of making stories as rich as possible without running out of time. It is also possible that they benefited from concrete knowledge about psychology during their course, which could have allowed them to integrate a more complex language in their descriptions of people and relationships at time two. The treatment group, on the other hand, could have regressed in the face of having confronted painful feelings in themselves during the experiential course (this possible eventuality and it's relationship to ego development and ego resilience will be discussed in greater depth later) or simply may have been weary of thinking in a more complex psychological manner about themselves and personal relationships, as had been required during the course. However, if the course had such an effect it was not reflected in defensiveness, which decreased for this group (non-significantly on the MMPI Dn scale, significantly on the MCSD scale). In terms of the unanticipated decrease in subjective mental health for the control group, it is possible that this was related to the fact that this group was preparing for their exam just before the posttest measures were administered. The treatment group had completed their final paper before post-test measures were administered. Thus external factors might have influenced an increase in anxious and

depressed feelings in the control group. Beyond such temporary external factors, it is unclear what would have instigated such a subjective decrease in well-being across members of the control group. Such increases in negative affects were not reflected in the Affect-Tone + Relationships scale of the SCORS, which did not change significantly between time 1 and time 2 in either group.

Finally, the limited time of the intervention and the interval at which outcomes were measured, may account for the unanticipated lack of change in self-report mental health or observer-rated mental health in the treatment group. Although limited interventions seem to have the capacity to change defensive functioning (Pennebaker & Beall, 1986; Pennebaker et al., 1988), it is quite likely that a 16 week intervention would not instigate large changes in ego functioning, given the finding that change in such enduring constructs may typically not occur until around the 2 year mark in treatment (Gordon, 2001). It is also likely that post treatment measures do not accurately reflect outcome, since several studies of dynamic treatments reflect a decrease in self-report and observer-rated functioning initially (around 3 to 6 months) that improves after longer intervals and can continue to improve post-treatment (Bateman & Fonagy, 1999; Bateman & Fonagy, 2000; Blomberg, Lazar, & Sandell, 2001).

Hypothesis 2b

It was expected that defensiveness would exert a suppressor effect on self-report mental health change between time 1 and time 2. This could occur if highly defensive individuals reported high mental health at time one, but equal or lower mental health at time two as a result of a decrease in defensiveness (and, therefore, increased awareness of or willingness to report pathology). A significant increase in self-report mental health

became apparent for the treatment group between time 1 and time 2 when the MMPI Dn scale was covaried (as expected), but not when the MCSD scale was covaried. It is unclear why covarying one defensiveness scale, but not the other, effected a significant increase in mental health between time one and time two for the treatment group, but it is likely an artifact of varying reliability and validity; recall that the MMPI Dn scale had a stronger relationship to observer-rated (positive) and self-report (negative) measures than the MCSD scale, and thus was more likely to be a significant covariate. However, this supportive finding does not necessarily demonstrate improvement in individuals who were defensive at time 1 because *observer* -rated mental health did not increase between time 1 and time 2 in the treatment group even with the MMPI-Dn scale covaried. Furthermore, although both measures of denial demonstrated mean decreases between time 1 and time 2 for the treatment group, the change was only significant for the MCSD scale. Thus, those who were not highly defensive at time 1 (based on the MMPI Dn scale) reported more improvement over the course of PSY 325. However, there is only marginal support for the suggestion that a *decrease* in defensiveness over the course of treatment accounts for the non-significant difference between pre and post subjective reports of mental health, because scores on the MMPI Dn scale did not decrease significantly. In short, there was some support for the possibility that defensiveness could attenuate results of treatment outcome studies that rely on self-report measures to demonstrate therapeutic change.

Characteristics and Possible Limitations of the SCORS-TAT

There are several indicators that characteristics of the SCORS-TAT may have lead to mixed findings in the present study. First, internal consistency of the SCORS

Affect-Tone + Relationships composite (12 items) was poor (alpha = .58). It is possible that card pull was responsible for much of the variability among participants' scores. Consistent with this possibility, the order of card presentation correlated significantly with scores on the measure. Order of presentation was controlled, but the variability this introduced may have weakened effects with the SCORS. However, eliminating cards that did not correlate well with other cards did not substantially improve the internal reliability among the cards. Coding early memories rather than TAT cards is an option often relied on in studies of this nature, but was not an option here because the treatment group worked explicitly with imagery of stressful early memories in the course, which would have confounded early memories data.

The positive correlation between Affect Tone + Relationships, self-report mental health, and defensiveness suggests that the SCORS was ineffective at detecting Pollyanna stories, despite explicit instructions to penalize stories that have unrealistic happy endings on the Affect Tone scale. If Pollyanna stories were detected, a negative correlation would have been expected at least between Affect tone + Relationships and the defensiveness measures. The inability of this composite scale to detect Pollyanna stories was unlikely due to errors in scoring as both coders scored practice data within acceptable reliability of the creator of the SCORS (ICC .81-.90), and scored the present data reliably with each other (ICC .82-.86).

The other composite SCORS measure (Complexity + Social Causality) also had notable shortcomings. Low self-report mental health was related to high Complexity + Social Causality even though one would have expected those who admitted to poor mental health to score poorly on an observer rated measure thereof as well. This could

have been due to the possibility that, in this non-clinical population, relatively low scores on the self-report mental health measures were more indicative of a forthcoming response to the measure than psychopathology per se. These observations together also suggested that SCORS scales may not be useful as global measures of mental health in isolation, or when scores are simply added together across scales.

A typological approach to the SCORS. In future research, taking better account of interactions *between* scales of the SCORS might be useful. For instance, past research has suggested that many individuals diagnosed with borderline personality disorder score high on complexity (describe people's intrapsychic lives in a rich and complex manner), but low on affect tone and relationships (make malevolent attributions about them; Ludolph et al., 1990; Westen et al., 1990) This line of thinking leads to the notion that combinations of high and low scores on the SCORS could be as unhealthy or less healthy than low scores *across* the SCORS subscales. The small size of the groups generated with this typological approach in the present study, prevented definitive conclusions but exploratory analyses were of interest.

Results of these analyses suggested some validity for the hypothesized SCORS groupings. Thus, the defensive/ Pollyanna group scored the highest on both defensiveness measures, whereas the group classified as borderline scored the lowest, although differences between the borderline, manifestly healthy, and manifestly unhealthy groups were non-significant.

Although SCORS scales have been both added together and analyzed separately in past research, past studied have not use the SCORS in a typological manner. This typological approach to the SCORS subscales highlights the likelihood that facets of ego

functioning that are considered adaptive in isolation (such as the ability to make complex attributions with regard to others' internal worlds) may be considered much less so when combined with poor functioning in other areas of ego development. Simply combining SCORS scales or analyzing them separately and relating specific scales to specific outcomes does not account for such potentially important typological effects. Future studies including a clinical population are necessary to elaborate the relationship between potential combinations of high and low functioning across SCORS subscales and self report mental health, defensiveness, and treatment outcomes.

Ego Strength and the SCORS. Readers may recall that the test-retest reliability of the SCORS was surprisingly low in this population (treatment group = .22, control group = .51). This author could find no report of test-retest reliability data on the SCORS in the literature. Others have also pointed out this lack (Lilienfield, et al., 2000) as a conspicuous gap in the literature that needs to be corrected.

Consistent with the poor test-retest reliability in this population is also the likelihood that ego strength may change rapidly, both in terms of progressions and regressions, in this older adolescent/young adult population.

A good deal of theoretical and empirical literature has been devoted to such variations. For instance, Erikson (1963) noted that changes in ego development are related to grappling with "psychosocial crises." Such crises represent, "a crucial point of increased vulnerability and heightened potential, and therefore the ontogenetic source of generational strength and maladjustment" (Erikson, 1963, p.294). Erikson understood this phenomenon as a result of confronting information that is inconsistent with an individual's current schema. Such confrontations precipitate what theorists often refer to

as dissonance, cognitive conflict, or disequilibrium (see Murray, 1983). In order to restore equilibrium, individuals may either attempt to assimilate information into current schemas or adapt one's schema to the new data, the later resulting in a transition to a higher level of ego development (Block, 1982; Loevinger, 1976). However, such transitions require relinquishing current schemas as well as current defenses or coping strategies, which often leads to dysphonic arousal and can result in temporary regressions or rapprochement. Perry (1968) recognized that late adolescence involves such pauses in development, lack of agency toward change, intellectual retreat, and escape in the form of dissociations. Rather than viewing such regressions as pathological, he described them as necessary for further growth.

It is unclear what pattern of results would reflect such a non-pathological regression in the present data. One could argue that the pattern of results found in the control group – an increase in self-reported symptomotology, but improvement in observer-rated mental health – could represent such a regression. However, it is unclear what would have systematically spurred on such a regression across the control group. Furthermore, it seems equally likely that such a regression would be reflected in lower scores on the Complexity + Social Causality composite, as it seems likely that in the face of such a regression, individuals would temporarily employ more ridged, less mature, defenses that would lead to more simplistic attributions about others' internal worlds and behaviors. There is mixed evidence for whether the intervention could have spurred on such a regression for individuals in the experimental group. Although the test-retest reliability of SCORS measures for this group is substantially lower than that of the control group (suggesting that the intervention had some effect on the SCORS) there was

no discernable or significant pattern to this change, and any non-pathological regression that may have occurred (which, in turn, could have attenuated initial improvement on the SCORS) was not evident in defensiveness measures, which did not increase for this group between time one and time two.

In light of the above possible interpretations of SCORS data, a further complication is the fact that it is unknown whether an individual's low SCORS score represents such a temporary regression or a more chronically low level of ego functioning (i.e. premature foreclosure in identity or ego development). Some authors have captured this continuum in terms of ego resilience -- the ability to fluidly adapt one's level of ego control based on the appropriateness of adaptation versus accommodation in a specific context (Block & Block, 1980). Thus a weakness in the present study was the inability of the measures used to account for ego resilience. The fact that younger age correlated with greater standard deviation on the SCORS is consistent with the possibility that such regressions and progressions may have had an important influence on these data.

This lack of accounting for ego resilience may also explain why the SCORS may not have been as useful an indicator of pathology in the present study as it has been in past studies using the scale. Most past studies using the SCORS have employed a psychiatric population or compared a psychiatric population to a "normal" group. Low ratings on the SCORS in a psychiatric population may be more likely to indicate a fixed developmental level, simply by virtue of a psychiatric diagnosis (particularly in the context of Axis II psychopathology, which the SCORS has frequently been used to assess). Thus, the inability of the scale to account for ego resilience was probably more detrimental in the present study (conducted with a population of college students) than

when used in the context of a psychiatric population. Although the SCORS has been used with "normal" college populations (Leigh, Westen, Barends, & Mendel, 1992) this limitation has not been discussed in previous literature on the measure.

CONCLUSIONS

The present study found mixed results with regard to the effects of defensiveness on self-report measures of mental health in treatment outcome research. Although defensiveness was related to self-reports of mental health in the expected manner (i.e. better self-reported mental health was related to higher defensiveness) defensiveness did not moderate the relationship between self-report mental health and observer-rated mental health. In short, self-report data may have been as good as observer-rated data in this study, even with defensiveness taken into account. Discrepancies between the two types of measures in this study may have simply reflected their quantification of separate facets of mental health (subjective awareness of pathology versus ego development). With regard to treatment outcome, defensiveness attenuated differences in self-report but not observer-rated mental health between time one and time two for the treatment group. Thus, there was some support for the hypothesis that taking defensiveness into account may yield more robust results in the treatment-outcome literature. The control group improved between time one and time two on one subscale (Complexity + Social Causality) and declined on self-report measures of mental health whereas the treatment group demonstrated no significant change in terms of the observer-ratings. It is unclear what accounts for this result but it could reflect non-pathological regression in the face of psychosocial crisis for either group, depending on how data are interpreted.

Future studies on the effects of excessive defensiveness on treatment outcome measurement should employ a psychiatric population and/or consider typologies. The present study employed college students enrolled in a course that is considered an analog to psychotherapy, but the observer-rated mental health measure likely generated scores

that were difficult to interpret in this population because the measure (the SCORS) does not account for ego resilience. As a result, it was not possible to distinguish between low scores that could have represented non-pathological regressions that are likely frequent in a college population and low scores that represented a more fixed and pathological level of ego development. If future studies are to use measures of ego development as a contrasting measure to self-report mental health, it may be more useful to employ measures that also account for ego resilience such as the California Q-sort (Block & Block, 1980). Future research should also focus on different effects of normal versus clinical levels of defensiveness, self-report mental health, and observer-rated mental health. The present study measured these as continuous variables for greater power rather than splitting data into "healthy" and "pathological" groups. Thus, it is unclear if clinical levels of defensiveness, observer-rated mental health, and self-report health may be related in fundamentally different ways than the present data are. Nonetheless, based on these data, there was scant support for the hypothesis that self-report measures are less accurate than observer-rated measures due to the effects of defensiveness on self-report.

REFERENCES

Adams, H. E., Wright, L. W., & Lohr, B. A. (1996). Is homophobia associated with sexual arousal? Journal of Abnormal Psychology, 105, 440-445.

Allen, J. P., Hauser, S. T., & Borman, E. (1996). Attachment theory as a framework for understanding sequelae of severe adolescent psychopathology: An 11-year follow-up study. Journal of Consulting and Clinical Psychology, 64, 254-263.

Bateman, A., & Fonagy, P. (1999). Effectiveness of partial hospitalization in the treatment of borderline personality disorder: A randomized clinical trial. *American Journal of Psychiatry, 156*(10), 1563-1569.

Bateman, A., & Fonagy, P. (2001). Treatment of borderline personality disorder with psychoanalytically oriented partial hospitalization: An 18-month follow-up. *American Journal of Psychiatry, 158*(1), 36-42.

Barends, A, Westen, D, Leigh, J., Silbert, D., Byers, S. (1990). Assessing Affect–Tone of Relationship Paradigms from TAT and interview data. Psychological Assessment, 2, 3, 329-332.

Bond, M. P., GarDenialer, S. T., Christian, J., & Sigal, J. J. (1983). Empirical study of self-rated defensive styles. Archives of General Psychiatry, 40, 333-338.

Beck, A. T., Steer, R. A., Brown, G. K. (1996) Beck Depression Inventory Manual (2nd Ed.). San Antonio, TX: The psychological Corporation.

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. Archives of General Psychiatry, 4 561-571.

Blatt, S. J., Brenneis, C. B., Schimek, J. G., & Glick, M. (1976). Normal development and psychopathological impairment of the concept of the object on the Rorschach. Journal of Abnormal Psychology, 85, 363-373.

Block, J. (1965). The Challenge of Response Sets. New York, NY: Appelton Century Crofts.

Blomberg, J., Lazar, A., & Sandell, R. (2001). Long-term outcome of long-term psychoanalytically oriented therapies: First findings of the Stockholm outcome of psychotherapy and psychoanalysis study. *Psychotherapy Research, 11*(4), 361-382.

Brenner, C. (1982) The Mind in Conflict. New York, NY: International Universities Press.

Hauser, S. T. & Safyer, A. W. The contribution of ego psychology to developmental psychopathology. In Cichetti, D & Cohen, D. J. (Eds.) Developmental

Psychopathology, Vol. 1: Theory and Methods. (pp. 555-580).

Cohen, J. (1992). A power primer. Psychological Bulletin, 112, 1 155-159.

Colvin, C. R., Block, J., & Funder, D. C. (1995). Overly positive self-evaluations and personality: Negative implications for mental health. Journal of Personality and Social Psychology, 68, 1152-1162.

Cooper, S. H. (1992). The empirical study of defensive processes: A review. In J. W. Barron, M. N. Eagle, & D. L. Wolitsky (Eds.), Interface of Psychoanalysis and Psychology (pp. 327-345).

Conklin, A. & Westen, D. (2002) Thematic apperception test. In Dorfman, W. & Hersan, M. (Eds.) Understanding psychological assessment: Perspectives on individual differences. (pp. 107-133) Dordrecht, Netherlands: Kluwer Academic Publishers.

Cramer, P. (1988). The Defense Mechanism Inventory: A review of the research and discussion of the scales. Journal of Personality Assessment, 52 (1), 142-164.

Cramer, P. (1989). The development of defense mechanisms. Journal of Personality, 55 (4), 597-614.

Cramer, P. (1991). Anger and the use of defense mechanisms in college students. Journal of Personality, 59, 39-55.

Crowne, D. P., & Marlowe, D. (1964). The approval motive: Studies in evaluative dependence. New York, NY: Wiley.

Davidson, K., & MacGregor, M. W. (1998). A critical aappraisal of self-report defense mechanism measures. Journal of Personality, 66 (6), 965-992.

Davis, P. J. (1987). Repression and the inaccessibility of affective memories. Journal of Personality and Social Psychology, 53, 585-593.

Dozier, M. & Kobak, R. R. (1992). Psychophysiology of attachment interviews: Converging evidence for deactivating strategies. Child Development, 63. 1473-1480.

Edwards, A. L. (1957). The Social Desirability Variable in Personality Assessment and Research. New York, NY: Dryden.

Endicott, J. E., Spitzer, R. L., Fleiss, J. L., & Cohen, J. (1976). The Global Assessment Scale: A procedure for measuring overall severity of psychiatric disturbance. Archives of General Psychiatry, 33, 766-771.

Erdelyi, M.H. (1990). Repression, reconstruction and defense. In J. Singer (Ed.)

Repression and dissociation: Implications for personality theory, psychopathology, and health (pp. 1-31). Chicago, Il. University of Chicago Press.

Exner, J. E. (1993). The Rorschach: A Comprehensive System. New York, NY. John Wiley and Sons.

Eysenck, H. J. (1994). Neuroticism and the illusion of mental health. American Psychologist, 49, 971-972.

Eysenck, H. J. and Eysenck, S. B. (1975). Eysenck Personality Questionnaire Manual. San Diego, CA. Educational and Industrial Testing Services.

Frank, L. K. (1948). Projective Methods. Springfield, Il.: Thomas.

Freud, A. (1936). The Ego and the Mechanisms of Defense. New York, NY: Wiley.

Freud, S. (1959/1895a) On the grounds for detaching a particular syndrome from neurasthenia under the description "anxiety neurosis." In J. Strachey (Ed. and Trans.), The standard edition of the complete psychological works of Sigmund Freud (Vol. 20, pp. 87-117). London: Hogarth Press.

Freud, S. (1959/1895b) The neuro-psychoses of defense. In J. Strachey (Ed. and Trans.), The standard edition of the complete psychological works of Sigmund Freud (Vol. 3, pp. 43-61). London: Hogarth Press.

Freud, S. (1959/1926) Inhibition, symptoms, and anxiety. In J. Strachey (Ed. and Trans.), The standard edition of the complete psychological works of Sigmund Freud (Vol. 20, pp. 75-175). London: Hogarth Press.

Gleser, G. C. and Iheilevich, D. (1969). An objective instrument for measuring defense mechanisms. Journal of Consulting and Clinical Psychology, 33, 51-60.

Gray, J. A. (1994). Framework for a taxonomy of psychiatric disorder. In S. H. Van

Goozen, N. E. Van de Poll, & J. A. Sergeant (Eds.) Essays on Emotion Theory (pp. 29-59). Hillsdale, NJ: Erlbaum.

Gordon, R. M. (2001). MMPI/MMPI-2 changes in long-term psychoanalytic psychotherapy. *Issues in Psychoanalytic Psychology, 23*(1-2), 59-79.

Gross, J. J. & Levenson, R. W. (1997). Hiding feelings: The acute effects of inhibition negative and positive emotion. Journal of Abnormal Psychology, 106 (1), 95-103.

Haan, N. (1963). Proposed model of ego functioning: Copingand defense mechanisms

in relationship to IQ change. Psychology Monographs, 77, 1-23.

Hibbard, S., Hisenroth, M. J., Hibbard, J. K., & Nash, M.R. (1995). A validity study of two projective object representation measures. Psychological Assessment, 7, 432-439.

Hamilton, M. (1960). A rating scale for depression. Journal of Neurology, Neurosurgery and Mental Science, 105, 985-987.

Hesse, E. (1999). The Adult Attachment Interview: Historical and current perspectives. In J. Cassidy and P. R. Shaver (Eds.), Handbook of attachment: theory, research and clinical applications. 395-433.

Holmbeck, G.N. (1997). Toward terminological, conceptual, and statistical clarity in the study of mediators and moderators: Examples from the child-clinical and pediatric psychology literatures. Journal of Consulting and Clinical Psychology, 65(4), 599-610.

Holmes, D. S. (1990). The evidence for repression: An examination of sixty years of research. In J. Singer (Ed.) Repression and dissociation: Implications for personality theory. Psychopathology, and health (pp. 85-102). Chicago, IL. University of Chicago Press.

Hughes, C. F., Uhlmann, C., & Pennebaker, J. W. (1994). The body's response to processing emotional trauma: Linking verbal text with autonomic activity. Journal of Personality, 62, 565, 585.

Jensen, M. R., (1987). Psychobiological factors predicting the course of breast cancer. Journal of Personality, 55 (2), 317-342.

Joiner, T. E., Schmidt, N. B., Lerew, D. R., Cook, J. H., Gencoz, T., Gencoz, F. (2000). Differential roles of depressive and anxious symptoms and gender in defensiveness. Journal of Personality Assessment, 75, 200-211.

Jorgenson, R. S., Schreer, G E., Baskin, L., & Kolodziej, M. (1992). Denail and the discrepancy between heart rate and reported negative affect: A study of convergent and discriminant validity. Psychotherapy and Psychosomatics, 58, 202-207.

Kaufman, G. (1996). The psychology of shame: Theory and treatment of shame based syndromes. New York, NY: Springer Publishing Company.

Kendall, P. C., Hollon, S. D., Beck, A. T., Hammen, C. L., Ingram, R. E. (1987). Issues and recommendations regarding the use of the Beck Depression Inventory. Cognitive Therapy and Research, 11, 289-299.

Krantz, D. & Manuck, S. (1984). Acute psychophysiologic reactivity and the risk of cartiovascular disease: A review and methodological critique. Psychological Bulletin, 96, 435-464.

Laplanche, J. & Pontalis, J. B. (1973). The Language of Psychoanalysis. New Yourk, NY: W. W. Norton & Company.

Lazarus, R. S. & Alfert, E. (1964). The short circuiting of threat by experimentally altering cognitive appraisal. Journal of Abnormal and Social Psychology, 69, 195-205.

Leigh, J., Westen, D., Barends, A., Mendel, M. (1992). The assessment of complexity of representations of people using TAT and interview data. Journal of Personality, 60, 4, 809-837.

Lilienfeld, S. O., Wood, J. M., & Garb, H. N. (2000). The scientific status of projective techniques. Psychological Science in the Public Interest,1 (2) 27-66.

Little, K. B. & Fisher, J. (1958). Two new experimental scales of the MMPI. Jounal of Consulting Psychology, 32, 305-306.

Loevinger, J. (1976). Ego Development. San Francisco, CA: Jossey-Bass.

Ludolph, P., Westen, D., Misle, B., Jackson, A., Wixom, J., & Wiss, F. C. (1990). The borderline diagnosis in adolescence: Symptoms and developmental history. American Journal of Psychiatry, 147, 470-476.

Manuck, S. B., Kaplan, J. R., & Clarkson, T. B. (1983). Behaviorally-induced heart rate reactivity and atherosclerosis in cynomolgus monkeys. Psychosomatic Medicine, 45, 95-108.

Mayman, M. (1968). Early memories and character structure. Journal of Projective Techniques and Personality Assessment, 32, 303-316.

McDermut, W., Miller, I., and Brown, R. (2001). The efficacy of group psychotherapy for depression: A meta-analysis and review of the empirical research. Clinical Psychology: Science and Practice 8,1 98-116.

McRoberts, C., Burlingame, G., Hoag, M. (1998). Comparative efficacy of individual and group psychotherapy: A meta-analytic perspective. Group Dynamics 2, 2 101-117.

Mikulincer, M. & Orbach, I. (1995). Attachment styles and repressive defensiveness: The accessibility and architecture of affective memories. Journal of Personality and Social Psychology, 68, 917-925.

Morokoff, P. J. (1985). Effects of sex guilt, repression, sexual "arousability," and sexual experience on female sexual arousal during erotica and fantasy. Journal of Personality and Social Psychology,49, 177-187.

Murray, H. (1943). Thematic Apperception Test Manual. Cambridge, MA: Harvard University Press.

Nolan-Hoeksema, S. (1987). Sex differences in unipolar depression: Evidence and theory. Psychological Bullitin, 101, 259-282.

Paulhus, D. L. (1991). Measurement and control of response bias. In Measurement of Personality and Social Psychological Attitudes J. P. Robinson, Shaver, P. R., & L. S. Wrightsman (Eds). New York, NY: Academic Press.

Perry, J. C. & Ianni, F. F. (1998). Observer-rated measures of defense mechanisms. Journal of Personality, 66 (6), 993-1024.

Perry, J. C., Cooper, S. H. (1989). An empirical study of defense mechanisms: Clinical Interview and life vignette ratings. Archives of General Psychiatry, 46, 444-460.

Pennebaker, J. W. (1987). Traumatic experiences and psychosomatic disease: Exploring the roles of behavioral inhibition, obsession, and confiding. Canadian Psychology, 26, 82-95.

Pennebaker, J. C. & Beall, S. (1986). Confronting a traumatic event: Toward an understanding of inhibition and disease. Journal of Abnormal Psychology, 95, 274-281.

Pennebaker, J. W., Glaser, J., & Glaser, R. (1988). Discloser of traumas and immune function: Health Implications for Psychotherapy. Journal of Consulting and Clinical Psychology, 56 (2), 239-245.

Pennebaker, J. W. & O'Heeron, R. C. (1984). Confiding in others and illness rates among spouses of suicide and accidental death victims. Journal of Abnormal Psychology, 93, 473-476.

Pennebaker, J. W. & Susman, J. R. (1988). Disclosure of traumas and psychosomatic processes. Social Science and Medicine, 26, 327-332.

Richards, J. M. & Gross, J. J. (2000). Emotion regulation and memory: The cognitive costs of keeping one's cool. Journal of Personality and Social Psychology, 79 (3), 410-424.

Scafidi, A., Field, T., Prodromidis, M., & Abrams, S. (1999). Association of fake-good MMPI-2 profiles with low Beck Depression Inventory scores. Adolescence, 34, 61-68.

Shedler, J., Mayman, M., & Manis, M. (1993). The illusion of mental health. American Psychologist, 48, 1117-1131.

Shedler, J., Mayman, M., and Manis, M. (1994). More Illusions. American Psychologist, 49, 974-976.

Spielberger, C. D. (1983) State-Trait Anxiety Inventory (form Y) Manual. Palo Alto, CA: Mind Garden.

Sterba, R. F. (1982). Reminiscences of a Viennese Psychoanalyst. Detroit, MI. Wayne State University Press.

Taylor, J. A. (1953). A personality scale of manifest anxiety. Journal of Abnormal and Social Psychology, 48, 285-290.

Taylor, S. E. & Brown, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. Psychological Bulletin, 103, 193-210.

Tillitski, C. (1990). A mta-analysis of estimated effect sizes for group versus indivudual bersus control treatments. International Journal of Group Psychotherapy, 40, 2 215-224.

Tinsley-Jones, H. (2003). Racism: Calling a spade a spade. Psychotherapy: Theory Research, Practice, Training, 40 (3), 179-186.

Tomkins, S. S. (1984). Affect theory. In K. R. Scherer and P. Ekman (Eds.), Approaches to Emotion Hillsdale, NJ: Erlbaum.

Vaillant, G. E. (1971). Theoretical hierarchy of adaptive ego mechanisms. Archives of General Psychiatry, 24, 107-118.

Vaillant, G. E. (1990). Repression in college men followed for half a century. In J. Singer (Ed.) Repression and dissociation: Implications for personality theory. Psychopathology, and health (pp. 259-273). Chicago, Il. University of Chicago Press.

Vaillant, G. E., & Drake, R. E. (1985). Maturity of ego defenses and relation to DSM-III Axis II Personality Disorder. Archives of General Psychiatry, 42, 597-601.

Watson, D. & Clark, L. A. (1984). Negative affectivity: The disposition to experience aversive emotional states. Psychological Bulletin, 96, 465-490.

Watson, C. G., Plemel, D., Vassar, P., Mainifold, V., Kucala, T., & Anderson, D. (1987). The

comparative validities of six MMPI repression scales. Journal of Clinical Psychology, 43, 472-477.

Weinberger, D. A., Schwartz, G. E., & Davidson, R. J. (1979). Low-anxious, high-anxious, and repressive coping styles: Psychometric patterns and behavioral and physiological responses to stress. Journal of Abnormal Psychology, 88 (4), 369-380.

Westen, D. (1998). Unconscious thought, feeling, and motivation: The end of a century-long debate. In Empirical Perspectives on the Psychoanalytic Unconscious. R. F. Bornstein & J. M. Masling (Eds.). Washington, DC: American Psychological Association.

Westen, D. (1991). Social cognition and object relations. Psychological Bulletin, 109, 429-455.

Westen, D., Klepser, J., Ruffins, S., Silverman, M., Lifton, N., & Boekamp, J. (1991). Object relations in childhood and adolescence: The development of working representations. Journal of Consulting and Clinical Psychology, 59, 3 400-409.

Weston, D., Lohr, N., Silk, K., Gold, L., & Kerber, K. (1990). Object relations and social cognitions in borderlines, major depressives, and normals: A Thematic Apperception test analysis. Psychological Assessment, 2, 355-364.

Westen, D., Lohr, N., Silk, K., Kerber, K., Goodrich, S. (1990). Social Cognition and Object Relations Scale (SCORS): Manual For Coding TAT Data. Ann Arbor: University of Michigan.

Westen, D., Ludolph, P., Block, M. J., Wixom, J., & Wiss, F. C. (1990). Developmental history and object relations in psychiatrically disturbed adolescent girls. American Journal of Psychiatry, 147, 1061-1068.

Wiesenthal, D. L. (1974). Some effects on the confirmation and disconfirmation of an unexpected monetary reward on compliance. Journal of Social Psychology, 92, 39-52.

APPENDIX A.

The following is a list of defenses used in this document that are not defined in the text. All are as defined by Laplanche and Pontalis (1967).

Anticipation: The fantasy rehearsal of future, potentially painful, events in an effort to achieve mastery over them or the affects involved.

Denial: See disavowal.

Disavowal: Refusing to recognize the reality of a traumatic perception.

Displacement: An idea's emphasis, interest, or intensity is detached from it and passed onto other ideas, which were originally of little intensity but which are related to the first idea by a chain of associations.

Identification with the aggressor: Faced with an external threat (typically represented by criticism emanating from an authority), the subject identifies himself with his aggressor. He may do so either by appropriating the aggression itself, or else by physical or moral emulation of the aggressor.

Intellectualization: Process whereby the subject, in order to master his conflicts and emotions, attempts to couch them in discursive form. The term denotes the preponderance of abstract thought over the emergence and acknowledgement of affects and fantasies.

Isolation (of affect): isolating thoughts or behavior so that their links with other thoughts or with the remainder of the subject's life are broken.

Negation: Procedure whereby the subject, while formulating one of his wishes, thoughts, or feelings, which has been repressed hitherto, contrives, by disowning it, to continue to defend himself against it.

Projection: Operation whereby qualities, feelings, wishes or objects, which the subject refuses to recognize or rejects in himself, are expelled from the self and located in another person or thing.

Rationalization: Procedure whereby the subject attempts to present an explanation that is either logically consistent or ethically acceptable for attitudes, actions, ideas, or feelings who's true motives are not perceived.

Reaction-formation: Psychological attitude diametrically opposed to a repressed wish, and constituted as a reaction against it (e.g. bashfulness counteracting exhibitionistic tendencies).

APPENDIX A.
(Continued)

Regression: The transition to modes of expression or though that are on a lower level as regards complexity, structure, and differentiation.

Repression: Process by which a subject attempts to repel, or to confine to the unconscious, representations which are bound to instinct.

Sublimation: The diversion of an instinctual aim toward a socially valued object.

Suppression: Mental operation of eliminating distressing or unwelcome contents. Suppression differs from repression in the sense that the procedure is of a conscious nature whereby the suppressed content passes only into the preconscious, but not into the unconscious.

APPENDIX B.

The following is a description and syllabus for PSY 325 – Affect and Self-Esteem written by Gershen Kaufman.

AFFECT AND SELF-ESTEEM

Development of psychological health through explorations of effect as the critical mediator of stress and self-esteem. Exploration of powerlessness-affect-stress cycles. Development of personal identity and interpersonal competence through translating psychological principles into effective tools for living.

I. Overview

What is most important to learn about living life is given least attention: How are we to create an inner sense of competence as well as effective power in the world? How are we to discover direction or purpose for ourselves? How are we to overcome life's stresses and manage its uncertainties? How can we learn to feel secure in an environment of accelerating uncertainty and powerlessness? How do we build satisfying and supportive relationships with peers, parents and partners? Problems in living arise when individuals have not learned how to effectively accomplish these essential developmental tasks. This course will offer an educational experience in developing these necessary psychological skills by exploring the sources of power and shame as foundations for competence. Through living consciously from a position of personal power, we learn to build a competent self.

II. Objectives

The specific objectives include reducing stress, enhancing self-esteem, developing personal identify, effectively managing affect, and developing interpersonal competence. Additionally, students will be taught a series of experiential tools which aid in building a competent self. These practical tools are translations of psychological principles into action.

III. Format

The course will combine weekly lectures with smaller section meetings to facilitate group discussion. The first class meeting of each week will be conducted as a lecture with all sections meeting together. The second class meeting of each week will be conducted as individual section meetings. Regular and consistent attendance is required.

IV. Required Texts

Kaufman, G. (1985). Shame: The Power of Caring. Cambridge: Schenkman.
Second Edition.

Kaufman, G., with Raphael, L. (1983). Dynamics of Power: Building a Competent
Self. Cambridge: Schenkman.

V. Written Assignments

1. Weekly Reaction Papers

A two-page reaction paper will be due each Thursday. These papers must (1) reflect your personal experience with the new tools presented each week and (2) demonstrate your knowledge of the theory and specific concepts presented in the assigned readings for that week. Papers must be typewritten, double-spaced, and handed in on time.

2. Term Paper

An eight-ten page integrative term paper will be due Monday of finals week. First, the term paper must demonstrate your mastery of course concepts by (1) relating the various psychological principles to their respective tools and (2) relating both principles and tools to the larger process of psychological health and self-esteem. You may discuss these, for example, in terms of the dynamics of stress and the dynamics of addiction. Second, describe in depth your experience applying concepts form this course in your personal life. Discuss your experience with the tools and concepts from an experiential as well as conceptual perspective. The paper must demonstrate your understanding of the theory behind the tools as well as how you have actually applied them in your own life. Finally, evaluate the effectiveness of your experiment in applying these principles and tools: Are they worth learning? Do they build personal competence, power, and competent relationships? You must include lectures, class discussions, personal experience with the tools, and both texts in organizing your paper. Be specific. Papers must be typewritten, double-spaced, and handed in on time.

VI. Grading

PSY 325 is a 3-credit, pass/no-grade course. In order to receive a P grade, students must complete all weekly reaction papers and term paper, and must participate in class meetings. Please note: Students are responsible for all material presented or work assigned when they are absent.

COURSE OUTLINE

Class Meetings	Units	Readings	Paper Due
Week 1 & 2	I. Powerlessness-Affect-Stress Power Powerlessness Affect Stress	Power, Ch. 1	Reaction Paper Due Each Thursday
Week 3 & 4	II. Shame and Self-Esteem Dynamics of Shame Shame Profile Self-Concept Addiction	Shame, Ch. 1-5	
Week 5 & 6	III. Identity: The Self's Relationship With the Self Inner Voices Guilt and Self-Care Inner Child Self-Affirming Identity	Power, Ch. 2	
Week 7 & 8	IV. Dimensions of Consciousness: Tools For Releasing Affect Owning Detachment Self-Observation Imagery	Power, Ch. 3	
Week 9 & 10	V. Interpersonal Competence Developmental Perspective Relationship Principles Relationship Process Power and Shame	Power, Ch. 4-5	Term Paper Due Monday of Finals Week

Reaction Paper #1

I. Tools (one page)

Discuss your personal experiences working with the following tools, give specific examples along with your honest reactions and also describe any difficulties or discoveries encountered:

1. Happiness List: Make a list of five events every day which leave you feeling happy, a smile on your face.

2. Adequacy List: Make a list of five events every day which leave you feeling proud of yourself.

II. Theory (one page)

Answer the following study questions based on Dynamics of Power, Ch. 1:

1. How do the happiness and adequacy tools relate to the four central concepts, responsibility, choice, living consciously and power.

2. How does power relate to powerlessness?

Reaction Paper #2

I. Tools (one page)

Discuss your personal experiences working with the following tools, give specific examples along with your honest reactions and also describe any difficulties or discoveries encountered:

1. Describe a current situation of powerlessness.

2. Identify your affective reactions during it.

3. Identify two choices for coping differently with that situation that could enable you to take back the power.

II. Theory (one page)

Answer the following study questions based on Dynamics of Power, Ch. 1:

1. What is the meaning of the concept of power as described in the text?

2. How does the need for power develop?

Reaction Paper #3

I. Tools (one page)

Discuss your personal experiences working with the following tools, give specific examples along with your honest reactions and also describe any difficulties or discoveries encountered:

1. Describe an old shame scene form childhood. It can be of any intensity or variety (embarrassment, shyness, guilt, self-consciousness, discouragement, inferiority). It can be from any setting (family, school, peer group). Describe your reactions during and after that scene.

2. Discuss how that old shame scene continues to affect you today, either positively or negatively.

II. Theory (one page)

Answer the following study questions based on Shame, Ch. 1-3:

1. What is the significance of shame and how does shame develop?

2. How does shame influence the development of personality and identity?

Reaction Paper #4

I. Tools (one page)

Discuss your personal experiences working with the following tools, give specific examples along with your honest reactions and also describe any difficulties or discoveries encountered:

1. Apply the Shame Profile to your own personality. Discuss the particular affects, needs, drives and purposes which, for you, have become fused with shame.

2. Discuss the role of shame in your own life, how shame has affected you.

II. Theory (one page)

Answer the following study questions based on Shame, Ch. 4, 5, and Epilogue:

1. What are the consequences and significance of shame internalization?

2. What are essential aspects of the healing or therapy process in regard to shame?

Reaction Paper #5

I. Tools (one page)

Discuss your personal experiences working with the following tools, give specific examples along with your honest reactions and also describe any difficulties or discoveries encountered:

1. Observe, accurately name and describe your inner voices, your characteristic negative identity scripts (self-blame, self-contempt, comparison-making). What is their source?

2. Attempt to replace negative voices/scripts with a new self-affirming voice/script and discuss your observations.

II. Theory (one page)

Answer the following study questions based on Dynamics of Power, Ch. 2:

1. How does the self's relationship to the self originate and what is necessary for changing it to a satisfying one?

2. What is the significance of defining "identity" as the "self's relationship with the self"?

Reaction Paper #6

I. Tools (one page)

Discuss your personal experiences working with the following tools, give specific examples along with your honest reactions and also describe any difficulties or discoveries encountered:

1. Experiment with inner child imagery or reparenting imagery and discuss your observations of the experience.

2. Work with the inner child tools through verbal or behavioral methods and discuss your observations.

II. Theory (one page)

Answer the following study questions based on Dynamics of Power, Ch. 2:

1. What is the meaning or significance of the inner child concept?

2. How does the inner child concept relate to the concept of identity?

Reaction Paper #7

I. Tools (one page)

Discuss your personal experiences working with the following tools, give specific examples along with your honest reactions and also describe any difficulties or discoveries encountered:

1. Practice the consultation with self tool daily and discuss your observations.

2. Discuss which affects, needs, and bodily states you are consciously aware of readily, and which are hardest for you to recognize.

3. Discuss scenes of purpose ("values" in the text) in terms of your imagined dreams for your future.

II. Theory

Answer the following study questions based on Dynamics of Power, Ch. 3:

1. Why is a "language of the self" necessary and useful?

2. How is "differentiated owning" related to "living consciously"?

Reaction Paper #8

I. Tools (one page)

Discuss your personal experiences working with the following tools, give specific examples along with your honest reactions and also describe any difficulties or discoveries encountered:

1. Practice several of the detachment tools (bubble meditation, refocusing attention, letting go imagery, detachment imagery, writing, humor, self-observation) and discuss you observations.

2. Discuss which methods are most effective for which situations.

II. Theory (one page)

Answer the following study questions based on Dynamics of Power, Ch. 3:

1. What is the significance of detachment and self-observation?

2. How does "imagery" relate to the other three dimensions of consciousness (owning, detachment and self-observation)?

Reaction Paper #9

I. Tools (one page)

Discuss your personal experiences working with the following tools, give specific examples along with your honest reactions and also describe any difficulties or discoveries encountered:

1. Discuss three different, current relationships (friend, family, romantic) along the following dimensions:

   a) Determine what you are needing, expecting, or looking for in each of the three relationships by observing your recurring relationship scenes.

   b) Objectively observe each of the three individuals in order to determine how well your expectations (imagined scenes) match reality in each relationship.

   c) Examine each of the three relationships for power/power-lessness and also for shame.

II. Theory (one page)

Answer the following study questions based on Dynamics of Power, Ch. 4 and 5:

1. What is the relationship between "staying defended" and "vulnerability"?

2. What is the relationship among power, shame and intimacy?

Term Paper

I. Tools (4-5 pages)

1. Describe in depth your experience applying concepts from this course in your personal life. Discuss your experience with the tools and concepts from an experiential as well as conceptual perspective.

2. The paper must demonstrate your understanding of the theory behind the tools as well as how you have actually applied them in your own life. Evaluate the effectiveness of your experiment in applying these principles and tools: Have they enhanced your own psychological health?

II. Theory (4-5 pages)

1. Demonstrate your mastery of course concepts by (1) relating the various psychological principles to their respective tools and (2) relating both principles and tools to the larger process of psychological health.

2. Discuss the dynamics of the four central dimensions of psychological health: self-esteem, stress, personal identity and interpersonal competence.

3. Discuss the dynamics of addiction and eating disorders.

APPENDIX C.

The Marlowe-Crown Social Desirability scale

Personal Reaction Inventory

Listed below are a number of statements concerning personal attitudes and traits. Read each item and decide whether the statement is true or false as it pertains to you.

1. Before voting I thoroughly investigate the qualifications of all the candidates.
2. I never hesitate to go out of my way to help someone in trouble.
3. It is sometimes hard for me to go on in my work f I am not encouraged.
4. I have never intensely disliked anyone.
5. On occasion, I have had doubts about my ability to succeed in life.
6. I sometimes feel resentful when I don't get my way.
7. I am always careful about my manner of dress.
8. My table manners at home are as good as when I eat out at a restaurant.
9. If I could get into a movie without paying and be sure I was not seen, I'd do it.
10. On a few occasions, I have given up doing something because I thought too little of my ability
11. I like to gossip at times
12. There have been times when I felt like rebelling against people in authority even though I knew they were right.
13. No matter who I'm talking to, I'm always a good listener.
14. I can remember "playing sick" to get out of something.
15. There have been occasions when I took advantage of someone.
16. I'm always willing to admit it when I make a mistake.
17. I always try to practice what I preach.
18. I don't find it particularly difficult to get along with loud-mouthed, obnoxious people.
19. I sometimes try to get even rather than forgive and forget
20. When I don't know something I don't at all mind admitting it.
21. I am always courteous, even to people who are disagreeable.
22. At times, I have really insisted on having things my own way.
23. There have been occasions when I felt like smashing things.
24. I would never think of letting someone else get punished for my wrongdoings.
25. I never resent being asked to return a favor.
26. I have never been annoyed when people express ideas different than my own.
27. I never make a long trip without checking the safety of my car.
28. There have been times when I was quite jealous of the good fortune of others.
29. I have almost never felt the urge to tell someone off.
30. I am sometimes irritated by people who ask favors of me.
31. I have never felt that I was punished without cause.
32. I sometimes think that when people have a misfortune, they only got what they deserved.
33. I have never deliberately said something that hurt someone's feelings.

APPENDIX D.

The Denial scale of the MMPI-2

1. I like to read newspaper articles on crime.
2. I enjoy detective or mystery stories
3. I feel that it is certainly best to keep my mouth shut when I'm in trouble
4. At times I feel like swearing.
5. I think a great many people exaggerate their misfortunes in order to gain sympathy and help of others.
6. It takes a lot of argument to convince most people of the truth.
7. I think most people would lie to get ahead.
8. Some people are so bossy that I feel like doing the opposite of what they request, even though I know they are right.
9. Most people will use somewhat unfair means to gain profit or advantage rather than lose it.
10. Often, I can't understand why I have been so irritable and grouchy.
11. I often wonder what hidden reason another person might have for doing something nice for me.
12. My conduct is largely controlled by the behavior of those around me.
13. I have often lost out on things because I couldn't make up my mind soon enough.
14. I resent having anyone trick me so cleverly that I have to admit I was fooled.
15. What others think of me does not bother me.
16. I frequently have to fight against showing that I am bashful.
17. I find it hard to make talk when I meet new people.
18. I wish I were not so shy.
19. In walking, I am very careful to step over sidewalk cracks.
20. I get mad easily and then get over it soon.
21. It is safer to trust nobody.
22. When in a group of people I have trouble thinking of the right things to talk about.
23. I drink an unusually large amount of water every day.
24. I am always disgusted with the law when a criminal is freed through the arguments of a smart lawyer
25. I am likely not to speak to people until they speak to me.
26. I can be friendly with people who do things which I consider wrong.

APPENDIX E.

The Beck Depression Inventory (BDI)

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

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

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

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

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

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

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

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

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

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

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

APPENDIX E (continued).
The Beck Depression Inventory

12. Worthlessness
    0 I do not feel I am worthless.
    1 I don't consider myself as worthwhile as I used to.
    2 I feel more worthless as compared to others.
    3 I feel utterly worthless.

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

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

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

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

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

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

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

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

APPENDIX F.

The State-Trait Anxiety Inventory

State Scale

Answers options for all items are (1) not at all, (2) somewhat, (3) Moderately so, or (4) very much so.

1. I feel calm.
2. I feel secure.
3. I am tense.
4. I feel strained.
5. I feel at ease.
6. I feel upset.
7. I am presently worried over possible misfortunes.
8. I feel satisfied.
9. I feel frightened.
10. I feel comfortable.
11. I feel saelf-confident.
12. I feel nervous.
13. I am jittery
14. I feel indecisive
15. I am relaxed.
16. I feel content.
17. I am worried.
18. I feel confused.
19. I feel steady.
20. I feel pleasant

Trait Scale

Answer options for all items are (1) almost never, (2) sometimes, (3) often, or (4) almost always.

21. I feel pleasant.
22. I feel nervous and restless.
23. I feel satisfied with myself.
24. I wish I could be as happy as others seem to be.
25. I feel like a failure.
26. I feel rested.
27. I am "calm cool, and collected.
28. I feel that difficultes are piling up so that I cannot overcome them.
29. I worry too much over something that doesn't really matter.
30. I am happy.
31. I have disturbing thoughts.
32. I lack self-confidence.
33. I feel secure.
34. I make decisions easily.
35. I feel inadequate.
36. I am content.
37. Some unimportant thought runs through my mind and bothers me.
38. I take disappointments so keenly that I can't put them out of my mind.
39. I am a steady person.
40. I get in a state of tension or turmoil as I think over my recent concerns and interests.

APPENDIX G

The supposition of a typological approach to the SCORS composites was initially suggested by the curve fit estimation that was conducted on the defensiveness scales and the SCORS scales. A significant inverse relationship was found and it appeared that a few individuals scoring extremely low on the MMPI-Dn scale, but very high on the Complexity + Social Causality scales of the SCORS, were responsible for this relationship. Visual inspection of the items endorsed by those scoring extremely low on the MMPI-Dn scale and the MCSD scale suggested these individuals might harbor a very cynical, perhaps paranoid world-view (quite opposite from the Pollyanna response style characterizing defensive responses on the two scales). Items endorsed by low scorers included the following: "(1) It is safer to trust no one" (2) "I think many people exaggerate their misfortunes in order to gain sympathy and help from others" (3) "I think most people would lie to get ahead" (4) "Most people will use somewhat unfair means to gain profit or advantage rather than lose it" (5) "There have been occasions when I took advantage of someone" (6) "I sometimes try to get even rather than forgive and forget." It seemed possible that many individuals who endorsed all or nearly all such items might fall within a borderline spectrum of interpersonal functioning. This would account for high Complexity scores as individuals diagnosed with borderline personality disorder frequently score high on this scale, but low on Affect Tone and Relationships scales (Westen et al., 1990a; Westen et al., 1990b).