The Core Self-Evaluations Scale: Development of a Measure

Timothy A. Judge and Amir Erez

Department of Management

University of Florida

Joyce E. Bono

Department of Psychology

University of Minnesota

Carl J. Thoresen

Department of Psychology

Tulane University

Author Note

Correspondence regarding this manuscript should be addressed to Timothy A. Judge,

Department of Management, Warrington College of Business, 211D Stuzin Hall, University of

Florida, Gainesville, FL 32611. Electronic mail should be sent to tjudge@ufl.edu.

Abstract

Despite an emerging body of research on a personality trait termed core self-evaluations, the trait continues to be measured indirectly. The present study reported the results of a series of studies that developed and tested the validity of the Core Self-Evaluations Scale (CSES), a direct and relatively brief measure of the trait. Results indicated that the 12-item CSES was reliable, displayed a unitary factor structure, correlated significantly with job satisfaction, job performance, and life satisfaction, and had validity equal to that of an optimal weighting of the four specific core traits (self-esteem, generalized self-efficacy, neuroticism, and locus of control), and incremental validity over the five-factor model. Overall, results suggest that the CSES is a valid measure that should prove useful in applied psychology research.

The Core Self-Evaluations Scale: Development of a Measure

A line of research has developed which suggests that a broad personality trait, termed core self-evaluations, is a significant predictor of job satisfaction and job performance. Judge, Locke, and Durham (1997) introduced the concept of core self-evaluations in an effort to provide a trait that would be a useful predictor job satisfaction, as well as perhaps other applied criteria. According to Judge et al. (1997), core self-evaluations is a broad, latent, higher-order trait indicated by four well-established traits in the personality literature: (1) self-esteem, the overall value that one places on oneself as a person (Harter, 1990); (2) generalized self-efficacy, an evaluation of how well one can perform across a variety of situations (Locke, McClear, & Knight, 1996); (3) neuroticism, the tendency to have a negativistic cognitive/explanatory style and to focus on negative aspects of the self (Watson, 2000); and (4) locus of control, beliefs about the causes of events in one's life—locus is internal when individuals see events as being contingent on their own behavior (Rotter, 1966). As one can gather from the commonality among these traits, core self-evaluations is a basic, fundamental appraisal of one's worthiness, effectiveness, and capability as a person.

The four core self-evaluations traits are some of the more prominent in psychology. Cumulatively, they have been referenced in more than 50,000 publications (PsycINFO search, October 20, 2001). Despite the salience of these traits, and some strong similarities between them, it has been relatively uncommon for researchers to study the traits together (Judge & Bono, 2001a). Even in the relatively rare case when the traits are studied together in personality research, generally they are treated as entirely separate variables with no discussion of their interrelationships or possible common core (e.g., Abouserie, 1994; Hojat, 1983; Horner, 1996). Similarly, in industrial-organizational (I-O) psychology research, various pairs of the core traits

have been related separately to a variety of outcomes, such as job performance (Bhagat & Chassie, 1978), career decision-making (Kishor, 1981), unemployment (Tiggemann & Winefield, 1984), or attributions (Hesketh, 1984). In a number of studies, however, Judge and colleagues have found that the four core traits load on a single factor (Erez & Judge, 2001; Judge, Erez, & Bono, 1998; Judge, Bono, & Locke, 2000; Judge, Locke, Durham, & Kluger, 1998), suggesting that it may be appropriate to consider the traits as indicators of a higher-order latent concept.

Beyond the empirical associations, support for the core self-evaluations concept can be derived from clear conceptual similarities among the traits. Pair by pair, the traits share conceptual similarities (see Judge & Bono, 2001a). It is our argument that the reason these surface traits share similarities is because they are indicators of a common core. Because core self-evaluations is a broad, latent trait that is the common source of the four (and perhaps other) specific traits, it is the psychological mechanism that causes these individual traits to be correlated. Because an individual who scores high on core self-evaluations is someone who is well-adjusted, positive, self-confident, efficacious, and believes in their own agency, it is this broad core that is then manifested in high levels of self-esteem, emotional stability, and general self-efficacy, and an internal locus of control. In short, we believe that high self-esteem and the other core traits result from a broad, general, positive self-regard.

By the same token, we do not contend that the four core traits are completely redundant. There may be parts of each that are unique and important. What we are arguing, though, is that there is considerable redundancy, and the latent concept of core self-evaluations explains this conceptual and empirical redundancy. Latent constructs exist at a deeper level than their indicators, and in fact causally influence the indicators or dimensions (Bollen & Lennox, 1991).

Thus, when an individual has a positive self-concept, measures of the four core traits are manifestations or indicators of this inner self-concept or core self-evaluation, and this explains why the four traits are conceptually and empirically related. Thus, rather than being a multidimensional aggregate construct, where a composite factor is comprised of dimensions that may or may not be related, core self-evaluations is a latent psychological construct because it is the "latent commonality underlying the dimensions" (Law, Wong, & Mobley, 1998, p. 747). In that way, self-esteem, generalized self-efficacy, and the other core traits are different ways in which core self-evaluations is realized.

Not only do the core traits appear to indicate a common factor, it appears that the concept is related to important work criteria. In two studies, research by Judge and colleagues has linked core self-evaluations to job satisfaction and further showed that intrinsic job characteristics mediated the relationship (Judge, Locke, et al., 1998; Judge et al., 2000). Judge et al. (2000) found that individuals with positive self-evaluations not only perceived their jobs as providing more intrinsic characteristics, they actually attained more challenging jobs (i.e., jobs with higher levels of job complexity). In a meta-analysis of 169 correlations, Judge and Bono (2001b) showed that the relation of four core traits to job satisfaction generalized across studies. In addition to meta-analyzing the relationship between core self-evaluations and job satisfaction, Judge and Bono (2001b) also investigated the correlation between the core traits and job performance. Analyzing 105 correlations, these authors found that the validity of three of the four core traits generalized across studies. The average validity was identical (ave. $\rho=.23$) to the validity of conscientiousness ($\rho=.23$; Barrick & Mount, 1991). In addition, Erez and Judge (2001) showed that core self-evaluations was related to motivation and performance in two separate studies. In the first study, these authors demonstrated that the core self-evaluations

factor was related to task motivation and performance in a laboratory setting. In the second study, they showed that the core trait was related to task activity, productivity as measured by sales volume, and the rated performance of insurance agents.

Despite impressive support for core self-evaluations, one issue that may interfere with future research is the measurement of the trait. In contemporary personality research, most traits are measured with relatively short, direct scales. For example, the best-known measures of conscientiousness measure the trait with scales that include nine (Benet-Martínez & John, 1998), 10 (Goldberg, 1999), or 12 (Costa & McCrae, 1992) items. In contrast, at present core self-evaluations is measured indirectly. In such studies (e.g., Judge, Locke, et al., 1998; Judge et al., 2000), the trait is not measured manifestly but is inferred to exist on the basis of its indicators. As a practical matter, this measurement strategy has several limitations.

First, the measures are indirect. This means that the core self-evaluations trait must be extracted by factor analyzing the four scales that indicate the trait (e.g., Judge, Erez, et al., 1998). A direct measure, because it is designed to precisely measure the underlying concept itself, rather than the indicators of the concept, may be more valid. The indirect measurement approach of past research also leads to confusion over whether the trait is a latent or aggregate construct (see below). Second, because of this indirect measurement from existing scales, the measure of core self-evaluations is relatively long. Judge, Locke, et al. (1998) and Judge et al. (2000) measured core self-evaluations with four scales that total 38 items. Given the relative brevity of measures of other traits, it would seem unnecessary to measure core self-evaluations with a combination of scales that, cumulatively, are relatively long. The length of the indirect measure may limit its usefulness, especially in organizational settings. Rather than utilizing a lengthy measure, some researchers may choose to measure only a single indicator (e.g., neuroticism or emotional

stability), and thereby miss a substantial amount of valid variance. A final possible limitation is that of empirical validity. The core traits display slightly differential relations with criterion variables (e.g., in Judge & Bono's [2001b] meta-analysis, emotional stability predicted the criteria less well than the other core traits, and the self-esteem-performance correlations were highly variable across studies); it is possible that a direct measure would achieve higher, and less variable, levels of validity.

Accordingly, the purpose of the present study is to report the results of a series of studies designed to develop and validate a direct measure of core self-evaluations, termed the Core Self-Evaluations Scale (CSES). Various criteria were used in evaluating the new measure. To test the validity of the measure, four independent samples were collected. In the next section of the paper, we discuss the nature of core self-evaluations in more detail, and then describe the criteria necessary to validate the new measure.

Relationship of Core Self-Evaluations to the Five-Factor Model of Personality

Because core self-evaluations is argued to be a broad trait in its own right, an important issue in the conceptualization of core self-evaluations is the relationship of this broad trait to the five-factor model of personality. In at least one way, core self-evaluations is closely linked to the five-factor model in that one trait—neuroticism—is common to both frameworks. This leads to two relevant questions: (1) does the concept of core self-evaluations have anything to offer beyond neuroticism, and (2) what is its relationship to other traits in the five-factor model? These questions bear directly on the usefulness of the core self-evaluations concept, and thus the utility of a direct measure of the concept. It is therefore important to deal with them.

Judge and Bono (2001a) have argued that typical measures of neuroticism are, both conceptually and operationally, so narrow that they fail to assess core aspects of the self. As

Hogan and Roberts (2001) have noted, most measures of personality originated from measures of

psychopathology, and nowhere is that more true than with neuroticism. The heart of most

measures of neuroticism is anxiety. Whereas anxiety can be seen as an important partial indicator

of neuroticism, core conceptions of self-worth are less prominent in measures of neuroticism.

One reason to believe that self-evaluation of one's worthiness or capabilities is underrepresented

in the five-factor model in general, and in measures of neuroticism in particular, is that in the

lexical approach to the five-factor model, evaluative adjectives were excluded from the adjective

list submitted to factor analysis. Thus, any conceptualization of neuroticism emerging from this

tradition would be narrower than the core self-evaluations construct in that one of the three

defining criteria for the core traits, evaluation-focus (Judge et al., 1997), is necessarily removed

from neuroticism. If this argument is correct, standard measures of neuroticism reflect anxiety

more than core evaluations of self-worth, and this may lead to the underprediction of important

criteria. For example, though measures of neuroticism display similar patterns of correlations

with job satisfaction and job performance compared to the other three core traits, of the four

traits, such measures are reliably the least valid correlates of both criteria (Judge & Bono,

2001b). Thus, it is possible that core self-evaluations is actually (low) neuroticism, but in such a

case neuroticism would need to be conceptualized and measured in a much broader manner.

Specifically, aspects of self-evaluation in general, and control and capabilities in particular,

would need to be included.

This leaves the issue of the relationship between core self-evaluations and two other Big

Five traits—conscientiousness and extraversion. On the one hand, core self-evaluations is

consistently related to extraversion and conscientiousness. Judge and Bono (2001a) reported that

neuroticism and self-esteem correlated significantly with extraversion (r=-.25 and r=.32,

respectively, $\underline{p} < .01$) and conscientiousness (r=-.39 and r=.39, respectively, $\underline{p} < .01$). These empirical associations make sense in that self-efficacy (belief in one's capability to perform) is sometimes seen as an aspect of conscientiousness (Costa & McCrae, 1992) and, of course, core self-evaluations. It is also reasonable to assume that those who tend to be more sociable, assertive, active, and upbeat (extroverts) will tend to evaluate themselves, their environment, and their control over their environment in a more positive manner. Thus, we expect core self-evaluations, and thus the CSES, to be moderately related to conscientiousness and extraversion.

On the other hand, evidence suggests that core self-evaluations is clearly distinct from the two traits. Factor analytic studies of the five-factor model suggest that one of the core traits—neuroticism—is distinct from conscientiousness and extraversion (Costa & McCrae, 1992). Since the other core traits correlate similarly with conscientiousness and extraversion as do neuroticism, a similar distinction should apply (Judge & Bono, 2001a). Furthermore, evidence indicates that the core self-evaluations independently contributes to criteria controlling for conscientiousness (e.g., Erez & Judge, 2001).

It is important to reiterate here that personality traits can be related, conceptually and empirically, without being redundant. There is good reason to believe that the CSES, like core self-evaluations, is related to conscientiousness and extraversion, but we do not believe that this necessarily leads to the conclusion that the CSES measures the same concept as conscientiousness or extraversion. This is, however, an assumption that can and will be tested.

Toward a Construct Valid Measure

Consistent with prior definitions of the core self-evaluations concept (Judge & Bono, 2001a; Judge et al., 1997; Judge, Locke, et al., 1998), several hypotheses, serving as necessary conditions for a favorable evaluation of the measure, were advanced. First, reliability is a

necessary but not sufficient condition for construct validity (Pedhauzer & Schmelkin, 1991). We estimate three forms of reliability: Internal consistency, test-retest (stability), and inter-rater reliability (Schwab, 1980). Second, if a measure is a reliable and valid assessment of the construct, the factor structure should match theoretical predictions (Schwab, 1980, p. 21). In this case, since the CSES is assumed to provide an overall measure of core self-evaluations, it should have a unitary factor structure. Third, another aspect of construct validation is to specify the nomological network (Cronbach & Meehl, 1955), wherein the relationship of the focal construct with other similar constructs is investigated. One component of the nomological network is the test of convergent validity. Convergent validity concerns the extent to which alternative measures of the construct share variance (Schwab, 1980). In this case, convergent validity can be estimated by the correlations of the four core traits with the CSES measure. If the measure is valid, it should display strong correlations with the four core traits.

Another component of the nomological network is discriminant validity. As Campbell (1960) noted, discriminant validity is "the requirement that a test not correlate too highly with measures from which it is supposed to differ" (p. 548). In the case of the CSES, discriminant validity could be evaluated by answering two critical questions: (1) despite its relationship with conscientiousness and extraversion, is the CSES distinct from these traits?; (2) is the CSES unrelated to other personality traits of the Big Five that theoretically seem to be completely distinct from core self-evaluations, such as agreeableness and openness? Thus, we expect the CSES to display significant correlations with extraversion and conscientiousness, and weak or nonsignificant correlations with agreeableness and openness. Furthermore, the magnitude of the correlations should be such that the CSES correlates more strongly with the core traits than with conscientiousness, extraversion, or the other Big Five traits.

Fourth, it is important to investigate the empirical validity of the measure. In the case of the CSES, we felt two of the most focal criteria in I-O psychology—job satisfaction and job performance—were appropriate criterion variables to investigate. In addition, we explored an important criterion in applied psychology more generally—life satisfaction or subjective well-being (Diener, Suh, Lucas, & Smith, 1999). We also examined whether CSES displays incremental validity in predicting these three applied criteria (job satisfaction, job performance, life satisfaction), when controlling for (a) the Big Five traits and (b) a factor comprised of each of the four core traits.

A fifth component of our validation efforts was to turn our attention to content validity (Schwab, 1980). Recognizing the importance of covering the broad domain of core self-evaluations in our measure, we imposed a requirement that the final scale include items reflecting the content of the core self-evaluations constructs. That is, the final scale must be true to our earlier definition of core self-evaluations and assess the source of the overlap among the four core traits.

In summary, we establish a series of necessary conditions to support the validity of the CSES. These conditions are that the CSES:

1. Is a reliable measure as assessed by (a) internal consistency, (b) test-retest reliability, and (c) inter-source (self-significant other) agreement.

2. Assesses a single dimensional construct.

3. Includes items that represent the construct of core self-evaluations, namely the source of the commonality among the four core traits.

4. Displays strong correlations with the four core traits: (a) self-esteem, (b) generalized self-efficacy, (c) locus of control, and (d) emotional stability.

5. Shows moderate correlations with conscientiousness and extraversion, and weak correlations with agreeableness and openness.

6. Shows empirical validity in correlating with central criteria in applied psychology, specifically (a) job satisfaction, (b) job performance, and (c) life satisfaction.

7. Shows usefulness in predicting the above criteria compared to the four core traits, (a) self-esteem, (b) generalized self-efficacy, (c) locus of control, and (d) emotional stability.

8. Shows incremental validity in predicting the above criteria controlling for other related, theoretically relevant traits, specifically conscientiousness and extraversion, and the five-factor model more generally.

Method

Samples

The data were collected from four independent samples—employees and managers of two companies and students at two different universities. In the field samples—Samples 1 and 2—we collected data on the CSES items, relevant personality traits (four core traits, several Big Five traits), and the outcome measures (job and life satisfaction, job performance). Both the field samples and the student samples allow us to investigate various aspects of the validity of the CSES. However, the student data also allowed us to test other aspects of the CSES, including the stability of the measure (Sample 3) and the inter-source reliability of the measure (Sample 4).

Participants in Sample 1 were employees from three locations of a midwestern food service company. Participants worked in jobs ranging from truck driver and warehouse employee to managers and sales representatives. Surveys were administered to employees on the job. Completed surveys were placed into sealed envelopes and returned to the authors in postage paid

envelopes. Participants were promised that their individual responses would remain confidential. Of the 365 employees in the organization, 280 completed surveys were returned, for a response rate of 77%.

Participants in Sample 2 were pharmaceutical salespersons employed by a large corporation headquartered in the eastern United States. A total of 256 sales representatives from the organization were surveyed. Survey materials were sent to the homes of all potential respondents. In total, 175 of the 256 employees returned usable survey packets, for a response rate of 68%. A majority (N=131, or 75%) of respondents were male. The mean age of respondents was 42.1 years (SD=11.0), and respondents reported being employed in their current positions for an average of 10.2 years.

Sample 3 consisted of undergraduates enrolled at a southeastern university. Participants completed a battery of personality and subjective well-being instruments in two class sessions over a one-month period of time. All participants received extra course credit in return for their participation. The same measures were collected at Time 1 and Time 2 (one month later). Two hundred and sixty-five individuals completed the first questionnaire and 205 completed the second questionnaire. Participant ages ranged from 18 to 40 years (M=20.4, SD=2.5); 54% were female.

Participants in Sample 4 were undergraduates at a large midwestern university. Extra credit points were offered in exchange for participation. There were 126 participants with an average age of 21.3 years; 57% of participants were male, and 86% were Caucasian. Participants completed the surveys in a classroom setting. Additionally, participants were asked to have a significant other (close friend or family member—somebody who knew them well) evaluate them using the CSES. In order to allow matching of self and other reports, participants were

asked to write a five-digit identification number on their own survey and on their significant

other survey. Significant others returned their surveys directly to the research team in postage

paid envelopes that were distributed with the surveys. Thus, participants did not have access to

the survey completed by their peer. Eighty-two percent of significant others reported knowing

the participant "very well."

Measures

     Neuroticism. Neuroticism was measured with the 12-item neuroticism scale from the

NEO-FFI Personality Inventory (Costa & McCrae, 1992) in the four samples. Example items

included, "I often feel inferior to others" and "Too often, when things go wrong, I get

discouraged and feel like giving up." In the present study, the coefficient alpha (α) reliabilities of

the scale were .87, .84, .89, and .89 for Samples 1-4, respectively.

     Self-esteem. Rosenberg's (1965) 10-item self-esteem scale was used to measure self-

esteem. This scale is the most common measure of self-esteem, and considerable empirical data

support its validity (Blascovich & Tomaka, 1991). Example items include, "I feel that I have a

number of good qualities," and "I feel that I am a person of worth, at least on an equal basis with

others." The reliabilities of the scale were α=.82, α=.80, α=.88, and α=.89 for Samples 1-4,

respectively.

     Generalized self-efficacy. Generalized self-efficacy was measured with a 10-item scale

developed by Judge, Locke, et al. (1998). Example items include, "When I make plans, I am

certain I can make them work" and "If something looks too complicated, I will not even bother

to try it." The reliabilities of the scale were α=.85, α=.80, α=.84, and α=.89 for Samples 1-4,

respectively.

Locus of control. Locus of control was measured with the internality sub-scale of Levenson's (1981) Internal, Powerful Others, and Chance (IPC) Scale. Levenson's measure exhibits moderate reliability and has been used in a wide variety of samples (Levenson, 1981). Example items include, "When I get what I want, it's usually because I worked hard for it," and "My life is determined by my own actions." The reliabilities of the scale were α=.57, α=.70, α=.69, and α=.60 for Samples 1-4, respectively.

Other Big Five traits. As was noted above, neuroticism was measured in all four samples. The Big Five traits of conscientiousness and extraversion also were measured in all four samples. The Big Five traits of agreeableness and openness were measured in Samples 2 and 4. The Big Five traits were measured with the NEO-FFI Personality Inventory (Costa & McCrae, 1992). The reliabilities of the conscientiousness scale were α=.87, α=.84, α=.88, and α=.90 for Samples 1-4, respectively. The reliabilities of the extraversion scale were α=.78, α=.79, α=.91, and α=.75 for Samples 1-4, respectively. The reliabilities of the agreeableness scale were α=.70, and α=.87 for Samples 2 and 4, respectively. The reliabilities of the openness scale were α=.73, and α=.68 for Samples 2 and 4, respectively.

Life satisfaction. Life satisfaction was measured in Samples 2, 3, and 4 with the Satisfaction With Life Scale (SWLS; Diener, Emmons, Larsen, & Griffin, 1985), a five-item measure of life satisfaction, consisting of such statements as, "In most ways my life is close to ideal." The reliability for this scale was α=.81 in Sample 2, α=.85 (time 1) and α=.83 (time 2) in Sample 3, and α=.85 in Sample 4.

Job satisfaction. Job satisfaction was measured in Samples 1 and 2 using five items from the Brayfield and Rothe (1951) measure of overall job satisfaction. The reliability and validity of this scale has been demonstrated in previous studies (e.g., Brooke, Russell, & Price, 1988).

Sample items are, "Most days I am enthusiastic about my job" and "I consider my job rather

unpleasant" (reverse scored). The reliabilities for this scale were $\alpha=.83$ and $\alpha=.82$ in Samples 1

and 2, respectively.

Job performance. In both Samples 1 and 2, job performance was measured using

supervisory ratings. The performance data were collected in Sample 1 one month after the

personality data were collected; in Sample 2, the personality and performance data were

collected concurrently (though obviously from different sources). In Sample 1, we used a broad

measure of job performance that included three task performance items (e.g., "In carrying out the

basic functions specified in his/her job description") and eight contextual performance items

(e.g., "In supporting and encouraging a coworker with a problem;" Motowidlo & Van Scotter,

1994), as well as one overall performance item ("On his/her overall job performance). These

items were rated on a 1=much below average to 5=much above average scale and were summed

to form an overall measure of job performance for each participant. In Sample 2, supervisors

rated sales representatives with respect to three global performance indices. The first evaluation

was a 15-item performance scale reflecting various aspects of salesperson performance. Sample

items included, "Overcomes difficult sales obstacles" and "Captures and holds market share for

products in a given class." Supervisors rated each salesperson on these items using a 0=Needs

much improvement to 8=Outstanding response scale. In addition, supervisors rated the quality,

quantity, and total performance using a three-item scale described by Judge, Thoresen, Pucik,

and Welbourne (1999), using the same 0-8 response scale. Finally, supervisors were asked to

provide an estimated percentile ranking (0% to 100%) for the overall performance of each

salesperson in response to the question, "Think of all the reps you currently supervise or have

supervised in the past—how does the overall job performance of this rep compare relative to all

the reps whom you currently supervise of have supervised in the past." These three measures

were standardized and summed to form a composite performance measure reflecting overall

salesperson job performance (standardized α=.96). In all cases, supervisors were instructed that

their ratings were for research purposes only and were completely confidential.

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

Stage 1: Item Development and Scale Construction

In constructing the Core Self-Evaluations Scale (CSES), we developed a pool of 65

items. These items were written based on a review of the literatures involving measurement of

the individual core traits (e.g., Chen, Gully, & Eden, 2001; Gray-Little, Williams, & Hancock,

1997; Levenson, 1981; Rotter, 1966; Rosenberg, 1965), as well as the literature on self-concept

(e.g., Baumeister, 1997). Furthermore, consistent with previous research in the personality (e.g.,

Goldberg, 1999) and applied psychology (e.g., Jackson, Wall, Martin, & Davids, 1993)

literatures, we also closely examined the items of the existing measures and where possible,

based our item development on these items. The items were generated to cover the range of core

self-evaluations as indicated by the four specific core traits. Accordingly, items were written that

involved evaluations of self-worth (e.g., "I wish I could have more self-respect"), evaluations of

one's control over one's environment (e.g., "I determine what will happen in my life"),

evaluations of one's capability and competence to be successful (e.g., "I am capable of coping

with most of my problems"), and evaluations of one's emotional adjustment (e.g., "There are

times when things look pretty bleak and hopeless to me"). Despite the origins of these items,

they were not written to be "pure" indicators of the individual core traits. To do so would be to

focus on the specific-trait variance attributable to each trait, rather than the general factor

variance that explains the associations among the traits. Thus, many of the items cut across the

four core traits such that some items may reflect a combination of two or more of the core traits. This was in keeping with the presumed nature of the core construct. As the above examples reveal, both positively-worded and negatively-worded items were generated.

Sample 1 was used to choose the scale items by applying several criteria. First, the items needed to adequately sample the content domain covered by the four individual core traits. Thus, the CSES measure had items that covered one or more of the core self-evaluations domains or, more often, the commonality among them.[1] Second, in order to ensure that the scale was reliable, items had to be significantly correlated with each other. Third, we examined the correlations of the items with job satisfaction, life satisfaction, and job performance; items that had unexpected or very weak correlations with these criteria were excluded. Finally, the scale had to be short enough to be useful. In all, we found 12 items that corresponded to these four criteria. Of the 12 items, six were positively-worded and six were negatively-worded (reverse-scored). These items are provided in Table 1. The properties of this scale were tested in the next three stages using all samples.

Stage 2: Psychometric Properties and Evidence of a General Factor

The purpose of this stage of the study was to explore the preliminary necessary conditions to establish the construct validity of the CSES. Several indicators of construct validity such as the reliability of the scale, its distributional properties, and its factor structure were investigated. Table 2 presents descriptive statistics on the CSES, as well as reliability estimates, across the data sets. As shown in Table 2, the distribution of the CSES was similar across the samples. The means ranged from 3.78 to 4.03 with an average of 3.87 and the standard deviations ranged from 0.47 to 0.59 with an average of 0.53. None of the means were significantly different from one another. Across the six measurements, all coefficient alpha

reliability estimates were above .80 with an average reliability of .84. Across the samples, the average item total correlations ranged from .48 to .55 with a mean of .50. Examination of the interitem correlation matrices in the six measurements revealed that all of the items were positively intercorrelated (once the negatively worded items had been reverse scored). The alphas, item-total correlations, and interitem correlations all suggest a high level of internal consistency. Test-retest reliability was .81, which shows good stability (test-retest reliability). The ICC value showed reasonable inter-rater reliability. Indeed, the value (ICC=.43) is exactly the same as the average self-peer ICC reported by Costa and McCrae (1992, Table 8) for the NEO PI-R. Together, these results show that the CSES is reliable and thus meets one necessary condition for the construct validity of the measure.

Confirmatory factor analysis, conducted using LISREL 8 (Jöreskog & Sörbom, 1996), was used to test the underlying structure of the scale. We tested three competing nested models using a variance-covariance matrix as input into the program. Since all specific core traits were implicitly represented in the scale, in these alternative models we loaded the 12 items on one of four specific factors we thought the item best (but not necessarily only) indicated (i.e., self-efficacy oriented items on a self-efficacy factor, self-esteem oriented items on a self-esteem factor, and so on). Thus, the alternative nested models consisted of four core factors. Because the CSES was created to tap the underlying core self-evaluation construct and not necessarily the distinct traits, the hypothesized model was that the 12 CSES items load on one higher order factor. This model was created by fixing all six relationships among the four factors to one. In the first alternative model, the four factors were allowed to freely correlate, and since all six pairwise correlations were freed, this model become a four factor model were the four constructs are hypothesized to be distinct.

Because previous research as well as the results of this study (see Table 4) show lower correlations between core self-evaluation and locus of control, in the second alternative model we tested whether the assumed locus of control items should be a part of the unitary scale. Therefore, we fixed all the relationship between the self-esteem, self-efficacy, and neuroticism factors to one but freed the relationship between the locus of control factor and the other three factors. As such, this model became a two-factor model were self-esteem, self-efficacy, and neuroticism are treated as part of a common second order factor whereas locus of control is treated as a separate factor. Because the CSES items were assessed on six different occasions across four samples (Time 1 and Time 2 in Sample 3 and self and significant other in Sample 4), six separate tests of these alternative models of the factor structure of the CSES were possible. In order to test the fit of these models, we report the following fit statistics: chi-square (χ^2) with corresponding degrees of freedom, Root-Mean-Square Residual (RMSR), Root-Mean-Square Error of Approximation (RMSEA), Goodness of Fit index (GFI), Comperative Fit Index (IFI), and Relative Fit Index (RFI). In each alternative model we also investigated the power of the test to reject the model if it was a misspecified model using LISPOWER. The fit statistics for the one factor model are reported in Table 3.

The average χ^2 of a single-factor model across the samples was 97.51 with 48 degrees of freedom (χ^2/df=2.03). The averages of the other fit statistics across the samples were: RMSR=.06; RMSEA=.08; GFI=.92; CFI=.92; RFI=.79. The power of the test to identify misspecified models ranged from .75 to 1.00. These fit statistics represented a good fit of the hypothesized model to the data across all samples. The power analysis results suggest that it is likely (ave. probability=.94) that the fit statistics would have led to a rejection of the hypothesized model if it were indeed false.

Table 3 shows that the fit statistics of the two alternative models across the six samples were very similar to the more parsimonious one factor model. The difference in chi-square tests show the four-factor model was not significantly different from the more parsimonious one factor model in any of the samples. When a multiple factor model shows similar fit to a single-factor model, the more parsimonious one-factor model is preferable (Bollen, 1989). Additionally, in five out of the six samples, the two-factor model (separating out the locus-oriented items) was not significantly different from the hypothesized one factor model. Thus, the overall results suggest that it is appropriate to view the CSES items as indicating a unidimensional factor.

Stage 3: Convergent and Discriminant Validity

Demonstrating the construct validity of a new scale requires assessing its convergent and discriminant validity. This can be accomplished by investigating whether the pattern of correlations of the focal concept with measures of other constructs conforms to theoretical propositions (Campbell, Trapnell, Heine, Katz, Lavallee, & Lehman, 1996). Therefore, in the third stage of the study, we investigated the nomological network of the CSES and examined the extent to which the pattern of correlations adhered to theoretical expectations.

Table 4 presents the correlations of the CSES with the four core traits and the Big Five dimensions. Both uncorrected correlations and correlations corrected for unreliability (based on internal consistency) are reported.[2] Since it has been claimed that core self-evaluations is a higher order construct that subsumes the four traits of self-esteem, self-efficacy, neuroticism, and locus of control, we expected the CSES to be substantially correlated with these four traits. Indeed, across the four samples, the CSES was highly correlated with self-esteem (average corrected correlation r_c=.87), generalized self efficacy (average r_c=.82), and neuroticism (average r_c=.76). Locus of control was less highly correlated with the CSES, with an average corrected

correlation of .50. Overall, though, the CSES showed significant convergence with the four core traits.

The correlations of the four remaining Big Five traits bears on the discriminant validity or distinctiveness of the CSES. We had expected the CSES to be significantly related to, but still distinct from, conscientiousness and extraversion, and essentially unrelated to openness and agreeableness. As shown in Table 4, both conscientiousness (average r_c=.51) and extraversion (average r_c=.50) were moderately to strongly correlated with CSES across the four samples. Furthermore, as shown in the table, the relationships of the CSES with agreeableness and openness were weak and non-significant. Overall, the patterns of relationships between CSES and the eight traits investigated adhered to theoretical expectations. Thus, although we have not yet developed a complete nomological network, the results suggest that the CSES is a valid construct inasmuch as it strongly converges with three of the four core traits, arguably diverges from other related traits (conscientiousness and extraversion), and clearly diverges from others (agreeableness and openness). Due to the correlations of the CSES with conscientiousness and extraversion, in a later section, we consider the incremental validity of the CSES over these traits.

Stage 4: Empirical Validity

We tested the empirical validity of the CSE scale using three criteria (i.e., job satisfaction, job performance, life satisfaction) that have been suggested in past research to be related to the core self-evaluation construct (Erez & Judge, 2001; Judge & Bono, 2001b; Judge et al., 2000; Judge, Locke, et al., 1998). The correlations between the CSES and the criterion variables in all the samples are reported in the bottom three rows of Table 4. A perusal of these correlations reveals that the new CSE measure was significantly related to the three criteria

across the samples—suggesting that the CSES has empirical validity. However, a new scale

should not only show good validity in terms of its correlations with various criteria, but also

demonstrate incremental validity beyond existing measures. In order to test the incremental

validity of the CSES, we conducted usefulness analysis (Darlington, 1990). We first tested the

contribution of the CSE scale over the prediction of the core self-evaluations factor. Consistent

with past research (see Erez & Judge, 2001, Judge et al., 1998), the core self-evaluation factor

was estimated by saving the results from a principal components analysis, where as expected

three traits loaded positively on a single factor and the fourth trait, neuroticism, loaded

negatively on the same factor. Thus, the purpose of the usefulness analysis was to investigate

whether there is an advantage to measuring core self-evaluation more directly with the CSES. In

this analysis, using hierarchical regression, the core self-evaluations factor was entered first to

predict the criterion variables, then the CSES was added to the equations to ascertain the increase

in the multiple correlation. These results were then compared to the reverse situation, whereby

the CSES was entered on the first step and the core factor was entered in the second step. As

shown in Table 5, usefulness analysis results indicated that, with a single exception (life

satisfaction in Sample 4), the addition of the CSES significantly increased the multiple

correlation across samples and criteria, whereas the core factor significantly increased the

multiple correlations in only two of seven cases.[3] Thus, these results suggest that the CSES better

predicts the criteria compared to the core factor as previously estimated by Judge and colleagues.

Second, we investigated to what extent one loses information by using the CSES instead

of the four original traits. In this analysis, using hierarchical regression, the four core traits were

entered first as a block to predict the criterion variables, then the CSES was added to the

equations to ascertain the increase in the multiple correlation. These results were then compared

to the reverse situation, whereby the CSES was entered on the first step and the four traits entered in the second step. The results show that in four of the seven equations, some information was lost using the CSES instead of the four core traits (because the core traits added significant variance beyond the CSES). However, in an equal number of regressions, information was also lost if the four traits and not the CSES were used (because the CSES added significant variance beyond the four core traits). Thus, this analysis indicates that the CSES performs as well as the four traits, and this of course does not consider the advantages provided by the single CSES variable versus four traits optimally weighted into a regression.

Finally, we also tested the incremental validity of the CSES beyond the contribution of the five-factor model. We should note that in several of the samples, agreeableness and openness were not measured. However, this should not be a problem because, as shown in Table 4, these two traits correlate weakly with the CSES. The Big Five traits were entered as a block to the usefulness analysis, either in the first or the second stage of the analysis in the same manner as described above. These results are presented in Table 5 and show that the CSES significantly contributed to six of the seven criteria beyond the contribution of the five-factor model. By the same token, the five-factor model contributed beyond the CSES in most cases as well. This suggests that each—the single CSES scale as well as the entire five-factor framework—add something unique. These values represent incremental validities and would be even more impressive if they were corrected for the biasing effects of measurement error. Considering these analyses as a whole, our findings indicate that the CSES contributed to the prediction of job and life satisfaction, and job performance, after controlling for a factor comprised of measures of the original core traits as well as the five-factor model traits, and in some cases beyond the contribution of the original four traits.

Discussion

We began this paper by arguing that a direct measure of core self-evaluations may have several advantages over the indirect measurement practices used in past research. Results from four independent samples supported the validity of the measure—the Core Self-Evaluations Scale (CSES)—whose development we described in this paper. Results suggest that the CSES is a useful means of assessing core self-evaluations. In terms of the validity of the CSES, our results suggest that the measure meets one condition for validity—it is reliable. The 12-item CSES displayed acceptable levels of internal consistency and test-retest reliability. Furthermore, the inter-source (self-significant other) level of agreement was comparable to that of other personality measures. For example, the self and peer reports for the CSES were correlated r=.43, a level of convergence similar to that typically obtained in research with established Big Five measures (Barbaranelli & Caprara, 2000; McCrae & Costa, 1987; Mount, Barrick, & Strauss, 1994). Results also suggested that the 12 CSES items loaded on a single dimensional construct.

Beyond reliability, several pieces of evidence support the validity of the CSES. First, it displayed convergent validity as evidence by its correlations with the four core traits. Second, it was significantly correlated with two of the focal criteria in industrial-organizational psychology—job satisfaction and job performance—in two independent samples. The CSES was significantly correlated with life satisfaction in three other samples. Third, in general, the CSES performed as well as an optimal weighting of the four core traits and displayed incremental validity in predicting these criteria controlling for the core self-evaluations factor as well as the traits from the five-factor model. We should note that these latter tests are quite rigorous as they pit a single 12-item measure against four individual traits in one analysis (the four core self-evaluations measures) and the five-factor model traits in another analysis. We acknowledge that

the entire five-factor model was not assessed in a couple of studies, though we seriously doubt this would influence the results as the excluded traits (agreeableness and openness) are those that correlate the least strongly with the CSES.

Why does the CSES generally predict the criteria better than, and even controlling for, the four individual core traits? After all, if the four core traits and the CSES measure the same construct—core self-evaluations—then why would the CSES predict better than the individual core traits? One explanation is that the CSES measures more than core self-evaluations. However, we have shown that the CSES does not appear to reflect conscientiousness or extraversion any more than do typical measures of the individual core traits. So, if the CSES does measure additional traits, those traits would have to lie outside of the five-factor framework. A more likely explanation is that the CSES more directly predicts criteria because it more directly measures the core self-evaluations construct. The CSES items were intended to tap the intersections among the core traits and it stands to reason that the CSES therefore measures this common core more directly, and with less error, than the individual core traits. In short, the CSES measures the commonality among the core traits, rather than the specific-factor variance attributable to the core traits themselves.

Despite substantial support for the validity of the CSES, there were several areas of concern, one of which involves its relatively strong correlation with conscientiousness. Across the four samples, on average, the CSES correlated .44 with conscientiousness. One might interpret these results as evidence that the CSES reflects, at least in part, conscientiousness. On the other hand, the correlation between the CSES and conscientiousness is weaker (.07, on average) than the relationship of a composite of the four core traits with conscientiousness. Thus, the CSES does not appear to reflect conscientiousness any more than do the four individual core

self-evaluations traits. Furthermore, though little discussed in the literature, emotional stability tends to correlate moderately strongly with conscientiousness (Judge & Bono, 2001a). For example, neuroticism and conscientiousness are correlated r=-.53 in the NEO normative database (Costa & McCrae, 1992). Thus, though the CSES does covary with conscientiousness, this relationship is considerably weaker than the correlations of the CSES with the individual core traits, and is no higher than the conscientiousness-neuroticism correlation more generally.

Another interesting aspect of the results is the moderate correlation between locus of control and the CSES. This has been an issue in the literature on core self-evaluations. For example, Judge et al. (2001b) reported a .62 factor loading of locus of control on the latent core self-evaluations, a value substantially lower than the average loading of the other three core traits (.90). Some of the explanation for the weaker association of locus of control is due to the lower reliability of locus of control. As Lefcourt (1981) has noted, measures of locus of control are notoriously unreliable. Indeed, when one corrects locus of control for unreliability in the present study, the average correlation with the CSES increases from .35 to .50. Nevertheless, locus of control, even when corrected for attenuation, is the core trait that correlates the weakest with the core self-evaluations in general and the CSES in particular. One might even argue that locus of control assesses conscientiousness as much as it does core self-evaluations. On the one hand, across the studies, locus of control correlated slightly more highly with conscientiousness (ave. r=.30) than it did with the other three core traits (ave. r=.27). On the other hand, the correlation of locus of control with the CSES was somewhat higher (ave. r=.35) than the average correlation with conscientiousness. Yet another complicating factor is that the locus-CSES correlation was no higher than the CSES-conscientiousness or CSES-extraversion correlations. Though these issues seemingly have more to do with the nature of core self-evaluations vis-à-vis its relation

with locus of control than with the validity of the CSES per se, whether locus of control belongs in core self-evaluations theory is an issue worthy of further research. Perhaps little is lost in dropping locus of control from the core self-evaluations taxonomy (and, by extension, perhaps locus evinces more incremental validity). Finally, it is conceivable that work-based measures of locus of control (Spector, 1988), as has been shown with self-esteem (Pierce, Gardner, Cummings, & Dunham, 1989), would yield more positive results.

A related issue concerns the relationship between core self-evaluations and emotional stability. Core self-evaluations, and thus the CSES, may simply be a measure of emotional stability. Because neuroticism is argued to be an indicator of core self-evaluations, and at the same time neuroticism is one of the most established traits in personality (Eysenck, 1990), it is relevant to ask whether core self-evaluations is simply another label for neuroticism. On the one hand, given neuroticism's solid foundation in personality psychology, and the strong correlations of the core traits with neuroticism, one might argue that it is more accurate and parsimonious to label core self-evaluations as neuroticism (Judge & Bono, 2001a).

On the other hand, if core self-evaluations and the measure developed here do measure emotional stability, this in itself is an advance. Judge and Bono (2001a) argue that many commonly used measures of neuroticism may simply be too narrow in scope to be as useful, especially in work contexts, as they might. From a content validity perspective, measures of neuroticism, perhaps owing to their psychopathological origins, tend to measure dysphoria, anxiety, and stress-proneness. As Hogan and Roberts (2001) recently noted, measures of psychopathology are overrepresented in the literature, yet are less relevant to predicting applied (i.e., work) criteria. By contrast, measures of neuroticism place considerably less emphasis on beliefs about one's capabilities and control. For example, there are no items in the neuroticism

scales of the NEO-FFI (Costa & McCrae, 1992b), the International Personality Item Pool (IPIP; Goldberg, 1999), or the Eysenck Personality Inventory (EPI; Eysenck & Eysenck, 1968) that explicitly reference control or capability. Thus, although core self-evaluations may be no broader than the theoretical concept of neuroticism, we believe that existing <u>measures</u> of neuroticism are too narrow to fully capture self-evaluations. Empirically, perhaps due to their origins, measures of neuroticism appear to be less valid predictors of work-related outcomes. Of the four core traits, neuroticism displays the weakest correlations with both job satisfaction and job performance (Judge & Bono, 2001b). By contrast, the CSES measure displays relatively strong correlations with job satisfaction and moderately strong correlations with job performance. In short, if the CSES measures (low) neuroticism, then it does so in a more empirically valid manner than do typical measures. Researchers interested in predicting applied criteria with a measure of emotional stability, then, would need to consider using the CSES or a comparably broad measure.

<u>Limitations</u>

Several limitations of the study should be noted. First, some might criticize the performance criteria on several grounds: (1) their subjectivity; (2) their concurrent nature; and (3) that a potential third variable such as ability may make the relationship spurious; (4) some of the criteria are self-report. Though we do not have direct evidence to dispute any of these criticisms here, there is ample reason to believe they are not serious problems.

With respect to the subjectivity of performance evaluations, this is the nature of virtually all performance criteria in organizational settings (e.g., a strong majority of the criteria in the Barrick & Mount [1991] personality-job performance meta-analysis were subjective). Objective performance data are generally only available for sales positions, which were not studied here.

More importantly, we are aware of no data suggesting that core self-evaluations uniquely predict subjective performance criteria better than objective criteria. In fact, there is evidence that the associations are similar (Erez & Judge, 2001). With respect to the concurrent nature of the study, it is possible that performance caused core self-evaluations, though, again, we are unaware of evidence on this point. Furthermore, the individual core trait of self-esteem appears to be as heritable (Roy, Neale, & Kendler, 1995) and stable (Conley, 1984) as neuroticism, and we are unaware of such claims being made with respect to neuroticism. Third, it is unlikely that ability is an important confounding variable here as evidence reliably indicates low correlations between personality and cognitive ability (McHenry, Hough, Toquam, & Hanson, 1990). Finally, with respect to the self-report nature of the criteria, it certainly is true the associations here may be inflated by response bias (assuming core self-evaluations is itself not a theoretically relevant source of response bias). However, we were careful to collect data that were not self-reported, including job performance in two studies, and life satisfaction in Study 4 was measured several months after the CSES was administered. Even under these conditions, the CSES still seemed to have empirical validity, suggesting it does not merely reflect the tendency to respond more positively to survey items (though we certainly agree this is part of what core self-evaluations is).

Applications and Future Research

We believe the CSES can be used in many applications in industrial-organizational and applied psychology. Most fundamentally, we believe that when individuals are interested in including any one of the four individual core self-evaluations traits in their studies, they should consider the possibility that the traits may be indicators of a broader concept. The results demonstrate that the overall measure of core self-evaluations outperforms any of the four individual measures in every criteria to which it was related in the four samples. This is not to

suggest that the CSES will better predict every criteria in every situation compared to the core traits. But, clearly, given the results presented here, researchers should recognize the possibility of under-prediction when using the four traits, particularly when the traits are used in isolation.

There are many criteria core self-evaluations can be used to predict. Evidence already indicates that, across studies, the core traits are related to job satisfaction and job performance (Judge & Bono, 2001b), and the CSES is related to these criteria as well. Given the underprediction afforded by neuroticism discussed earlier, it is particularly important that selection decision-makers utilize a broad measure such as the CSES rather than a traditional measure of neuroticism. There are other criteria that the individual core traits have been found to predict, such as coping with organizational change (Judge et al., 1999), organizational justice (Vermunt, van Knippenberg, van Knippenberg, & Blaauw, 2001), voice behavior (LePine & Van Dyne, 1998), among others. These are outcomes organizations likely value, and given the results observed here, it is likely that the validities observed in previous studies would be even higher if the CSES were used.

From the standpoint of basic psychological research, there is a need for future research examining the antecedents of core self-evaluations. This work might focus on determining the extent to which core self-evaluations are stable traits with a genetic basis or whether they are malleable, subject to change based on performance or even life events. Future research on the relationship of locus of control to core self-evaluations also is needed. As is true with the development of any new scale, there is a need for ongoing validation of the CSES. This measure should be tested in a variety of settings, including those where its predictive validity can be further assessed.

References

Abouserie R. (1994). Sources and levels of stress in relation to locus of control and self esteem in university students. Educational Psychology, 14, 323-330.

Arthur W Jr, Doverspike D, Barrett GV. (1996). Development of a job analysis-based procedure for weighting and combining content-related tests into a single test battery score. Personnel Psychology, 49, 971-985.

Barbaranelli C, Caprara GV. (2000). Measuring the Big Five in self-report and other ratings: A multitrait-multimethod study. European Journal of Psychological Assessment, 16, 31-43.

Barrick MR, Mount MK (1991). The Big Five personality dimensions and job performance: A meta-analysis. Personnel Psychology, 44, 1-25.

Baumeister RF. (1997). Identity, self-concept, and self-esteem: The self lost and found. In R Hogan, JA Johnson, & S Briggs (Eds.), Handbook of personality psychology (pp. 681-710). San Diego, CA: Academic Press.

Benet-Martínez V, John OP. (1998). Los Cinco Grandes across cultures and ethnic groups: Multitrait-multimethod analyses of the Big Five in Spanish and English. Journal of Personality and Social Psychology, 75, 729-750.

Bhagat RS, Chassie MB. (1978). The role of self-esteem and locus of control in the differential prediction of performance, program satisfaction, and life satisfaction in an educational organization. Journal of Vocational Behavior, 13, 317-326.

Blascovich J, Tomaka J. (1991). Measures of self-esteem. In JP Robinson, PR Shaver, & LS Wrightsman (Eds.), Measures of personality and social psychological attitudes (Vol. 1, pp. 115-160). San Diego, CA: Academic Press.

Bollen KA. (1989). Structural equations with latent variables. New York: Wiley.

Brayfield AH, Rothe HF. (1951). An index of job satisfaction. Journal of Applied Psychology, 35, 307-311.

Brooke PP, Russell DW, Price JL. (1988). Discriminant validation of measures of job satisfaction, job involvement, and organizational commitment. Journal of Applied Psychology, 73, 139-145.

Campbell DT. (1960). Recommendations for APA test standards regarding construct, trait, or discriminant validity. American Psychologist, 15, 546-553.

Campbell JD, Trapnell PD, Heine SJ, Katz IM, Lavallee LF, Lehman DR. (1996). Self-concept clarity: Measurement, personality correlates, and cultural boundaries. Journal of Personality and Social Psychology, 70, 141-156.

Chen G, Gully SM, Eden D. (2001). Validation of a new generalized self-efficacy scale. Organizational Research Methods, 4, 62-83.

Cohen J, Cohen P. (1983). Applied multiple regression/correlation analysis for the behavioral sciences (2nd ed.). Hillsdale, NJ: Erlbaum.

Conley JJ. (1984). The hierarchy of consistency: A review and model of longitudinal findings on adult individual differences in intelligence, personality and self-opinion. Personality and Individual Differences, 5, 11-25.

Costa PT Jr, McCrae R. (1992). NEO-PI-R and NEO-FFI professional manual. Odessa, FL: Psychological Assessment Resources.

Cronbach LJ, Meehl PE. (1955). Construct validity in psychological tests. Psychological Bulletin, 52, 281-302.

Darlington RB. (1990). Regression and linear models. New York: McGraw-Hill.

Day DV, Silverman SB. (1989). Personality and job performance: Evidence of incremental validity. Personnel Psychology, 42, 25-36.

Diener E, Emmons RA, Larsen RJ, Griffin S. (1985). The Satisfaction With Life Scale. Journal of Personality Assessment, 49, 71-75.

Diener E, Suh EM, Lucas RE, Smith HL. (1999). Subjective well-being: Three decades of progress. Psychological Bulletin, 125, 276-302.

Erez A, Judge TA. (2001). Relationship of core self-evaluations to goal setting, motivation, and performance. Journal of Applied Psychology, 86, 1270-1279.

Eysenck HJ. (1990). Biological dimensions of personality. In LA Pervin (Ed.), Handbook of personality: Theory and research (pp. 244-276). New York: Guilford Press.

Eysenck HJ, Eysenck SBG. (1968). Manual of the Eysenck Personality Inventory. San Diego, CA: Educational and Industrial Testing Service.

Goldberg LR. (1999). A broad-bandwidth, public-domain, personality inventory measuring the lower-level facets of several five-factor models. In I Mervielde, IJ Deary, F De Fruyt, & F Ostendorf (Eds.), Personality psychology in Europe (Vol. 7, pp. 7-28). Tilburg, The Netherlands: Tilburg University Press.

Gray-Little B, Williams VSL, Hancock TD. (1997). An item response theory analysis of the Rosenberg Self-Esteem Scale. Personality and Social Psychology Bulletin, 23, 443-451.

Guion RM. (1991). Personnel assessment, selection, and placement. In MD Dunnette & LM Hough (Eds.), Handbook of industrial and organizational psychology (2nd ed., vol. 2, pp. 327-397). Palo Alto, CA: Consulting Psychologists Press.

Harter S. (1990). Causes, correlates, and the functional role of global self-worth: A life span perspective. In RJ Sternberg & J Kolligan Jr (Eds.), Competence considered (pp. 67-97). New Haven: Yale University Press.

Hesketh B. (1984). Attribution theory and unemployment: Kelley's covariation model, self-esteem, and locus of control. Journal of Vocational Behavior, 24, 94-109.

Hogan RT, Roberts BW. (2001). Introduction: Personality and industrial and organizational psychology. In BW Roberts & R Hogan (Eds.), Personality psychology in the workplace (pp. 3-16). Washington, DC: American Psychological Association.

Hojat M. (1983). Comparison of transitory and chronic loners on selected personality variables. British Journal of Psychology, 74, 199-202.

Horner KL. (1996). Locus of control, neuroticism, and stressors: Combined influences on reported physical illness. Personality and Individual Differences, 21, 195-204.

Jackson PR, Wall TD, Martin R, Davids K. (1993). New measures of job control, cognitive demand, and production responsibility. Journal of Applied Psychology, 78, 753-762.

Jöreskog K, Sörbom D. (1996). LISREL 8: User's reference guide. Chicago: Scientific Software International.

Judge TA, Bono JE. (2001a). A rose by any other name…Are self-esteem, generalized self-efficacy, neuroticism, and locus of control indicators of a common construct? In BW Roberts & R Hogan (Eds.), Personality psychology in the workplace (pp. 93-118). Washington, DC: American Psychological Association.

Judge TA, Bono JE. (2001b). Relationship of core self-evaluations traits—self-esteem, generalized self-efficacy, locus of control, and emotional stability—with job satisfaction and job performance: A meta-analysis. Journal of Applied Psychology, 86, 80-92.

Judge TA, Bono JE, Locke EA. (2000). Personality and job satisfaction: The mediating role of job characteristics. Journal of Applied Psychology, 85, 237-249.

Judge TA, Erez A, Bono JE. (1998). The power of being positive: The relationship between positive self-concept and job performance. Human Performance, 11, 167-187.

Judge TA, Locke EA, Durham CC. (1997). The dispositional causes of job satisfaction: A core evaluations approach. Research in Organizational Behavior, 19, 151-188.

Judge TA, Locke EA, Durham CC, Kluger AN. (1998). Dispositional effects on job and life satisfaction: The role of core evaluations. Journal of Applied Psychology, 83, 17-34.

Judge TA, Thoresen CJ, Pucik V, Welbourne TM. (1999). Managerial coping with organizational change: A dispositional perspective. Journal of Applied Psychology, 84, 107-122.

Kishor N. (1981). The effect of self-esteem and locus of control in career decision making of adolescents in Fiji. Journal of Vocational Behavior, 19, 227-232.

Law KS, Wong CS, Mobley WH. (1998). Toward a taxonomy of multidimensional constructs. Academy of Management Review, 23, 741-755.

LePine JA, Van Dyne L. (1998). Predicting voice behavior in work groups. Journal of Applied Psychology, 83, 853-868.

Levenson H. (1981). Differentiating among internality, powerful others, and chance. In HM Lefcourt (Ed.), Research with the locus of control construct (pp. 15-63). New York: Academic Press.

Locke EA, McClear K, Knight D. (1996). Self-esteem and work. International Review of Industrial/Organizational Psychology, 11, 1-32.

McCrae RR, Costa PT Jr. (1987). Validation of the five-factor model across instruments and observers. Journal of Personality and Social Psychology, 52, 81-90.

McHenry JJ, Hough LM, Toquam JL, Hanson MA. (1990). Project A validity results: The relationship between predictor and criterion domains. Personnel Psychology, 43, 335-354.

Motowidlo SJ, Van Scotter JR. (1994). Evidence that task performance should be distinquished from contextual performance. Journal of Applied Psychology, 79, 475-480.

Mount MK, Barrick MR, Strauss JP. (1994). Validity of observer ratings of the Big Five personality factors. Journal of Applied Psychology, 79, 272-280.

Pedhazur EJ, Schmelkin LP. (1991). Measurement, design, and analysis: An integrated approach. Hillsdale, NJ: Erlbaum.

Pierce JL, Gardner DG, Cummings LL, Dunham RB. (1989). Organization-based self-esteem: Construct definition, measurement, and validation. Academy of Management Journal, 32, 622-648.

Rosenberg M. (1965). Society and the adolescent self-image. Princeton, NJ: Princeton University Press.

Rotter JB. (1966). Generalized expectancies for internal versus external control of reinforcement. Psychological Monographs, 80 (1, Whole No. 609).

Roy MA, Neale MC, Kendler KS. (1995). The genetic epidemiology of self-esteem. British Journal of Psychiatry, 166, 813-820.

Schwab DP. (1980). Construct validity in organizational behavior. Research in Organizational Behavior, 2, 3-43.

Spector PE. (1988). Development of the Work Locus of Control Scale. Journal of Occupational Psychology, 61, 335-340.

Tiggemann M, Winefield AH. (1984). The effects of unemployment on the mood, self-esteem, locus of control, and depressive affect of school-leavers. Journal of Occupational Psychology, 57, 33-42.

Vermunt R, van Knippenberg D, van Knippenberg B, Blaauw E. (2001). Self-esteem and outcome fairness: Differential importance of procedural and outcome considerations. Journal of Applied Psychology, 86, 621-628.

Viswesvaran C, Ones DS, Schmidt FL. (1996). Comparative analysis of the reliability of job performance ratings. Journal of Applied Psychology, 81, 557-574.

Watson D. (2000). Mood and temperament. New York: Guilford Press.

Footnotes

[1] Though selection of the CSES items was inspired by an understanding of the four individual core traits, and knowledge of existing measures, the items did not strictly belong to the measures. Indeed, many of the items could be argued to sample multiple trait domains. For example, the item "I often feel helpless" might be argued to reflect any of the four core traits. Given that the motivation was to capture the commonality among the individual traits, this was seen as an appropriate, and even desirable, situation.

[2] Unreliability in job performance was corrected based on coefficient alpha. To make the CSES-job performance correlations comparable to meta-analytic estimates reported in the literature (e.g., Barrick & Mount, 1991), one would need to correct job performance for unreliability using estimates of inter-rater reliability. Since estimates of the inter-rater reliability of performance ratings are lower than estimates of internal consistency reliability (Viswesvaran, Ones, & Schmidt, 1996), the CSES correlations corrected for inter-rater reliability would be higher than those reported in Table 4.

[3] The statistical formula (Cohen & Cohen [1983], p. 145) for testing the significance of the change in the R^2 is the following:

$$F = \frac{(R_F^2 - R_R^2)/K_F}{(1 - R_F^2)/(n - K_R - K_F - 1)}$$

Where:

$R_F^2 = R^2$ of the full model; $R_R^2 = R^2$ of the reduced model; K_F=number of explanatory variables in the full model; K_R=number of explanatory variables in the reduced model. Due to the nature of the formula, the same ΔR^2 values can have differential levels of significance depending on the other parameters in the formula.

Table 1

The Core Self-Evaluations Scale (CSES)

Instructions: Below are several statements about you with which you may agree or disagree. Using the response scale below, indicate your agreement or disagreement with each item by placing the appropriate number on the line preceding that item.

1	2	3	4	5
Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree

1. _____ I am confident I get the success I deserve in life.

2. _____ Sometimes I feel depressed. (r)

3. _____ When I try, I generally succeed.

4. _____ Sometimes when I fail I feel worthless. (r)

5. _____ I complete tasks successfully.

6. _____ Sometimes, I do not feel in control of my work. (r)

7. _____ Overall, I am satisfied with myself.

8. _____ I am filled with doubts about my competence. (r)

9. _____ I determine what will happen in my life.

10. _____ I do not feel in control of my success in my career. (r)

11. _____ I am capable of coping with most of my problems.

12. _____ There are times when things look pretty bleak and hopeless to me. (r)

Notes: r=reverse-scored. This measure is non-proprietary (free) and may be used without permission.

Table 2

Distributional Properties and Reliability Estimates of the Core Self-Evaluations Scale (CSES)

	Mean (SD)	Internal consistency reliability	Test-retest reliability	Inter-source reliability
		Distributional Properties and Reliability Estimates		
Sample 1	3.83 (.48)	.85		
Sample 2	4.03 (.58)	.83		
Sample 3			.81	
Time 1	3.83 (.57)	.85		
Time 2	3.95 (.59)	.87		
Sample 4				.43
Self	3.78 (.50)	.83		
Other	3.85 (.47)	.81		

Notes: Other=significant other. Internal consistency estimates are coefficient alpha (α) reliability estimates. Test-retest estimates are correlations at Time 1 and Time 2. Inter-source estimate is the correlation between self and significant other reports as estimated by the Intraclass Correlation (ICC).

Table 3

Fit Statistics from Confirmatory Factor Analysis of Single Dimensional Structure of Core Self-Evaluations Scale (CSES)

Fit statistic	Sample 1 (N=279)			Sample 2 (N=279)			Sample 3 (N=261)		
Number of factors in model	One	Four	Two	One	Four	Two	One	Four	Two
Degrees of freedom	48	42	45	48	42	45	48	42	45
Chi-square (χ^2)	140.40	132.45	137.55	87.95	80.15	84.43	104.84	93.58	95.97
Difference in chi-square ($\Delta\chi^2$)	---	7.95	2.85	---	7.80	3.52	---	11.26	8.87*
Root-mean-square error	.06	.06	.06	.07	.07	.06	.05	.05	.05
RMSEA	.09	.09	.09	.08	.08	.08	.06	.07	.07
Goodness of fit index	.92	.92	.92	.92	.92	.92	.94	.94	.94
Comparative fit index	.92	.92	.92	.93	.93	.93	.93	.94	.94
Relative fit index	.84	.82	.83	.80	.79	.80	.84	.84	.84
Power at .05 significance level	1.00	1.00	1.00	.97	.97	.97	.98	.98	.99

Table 3 Continues

Table 3 (Continued)

Fit Statistics from Confirmatory Factor Analysis of Single Dimensional Structure of Core Self-Evaluations Scale (CSES)

Fit statistic	Sample 3-Time 2 (N=205)			Sample 4-Self (N=126)			Sample 4-Other (N=126)		
Number of factors in model	One	Four	Two	One	Four	Two	One	Four	Two
Degrees of freedom	42	48	45	48	42	45	48	42	45
Chi-square (χ^2)	84.65	78.12	82.89	98.29	85.81	94.54	75.44	63.14	68.74
Difference in chi-square ($\Delta\chi^2$)	---	6.53	1.76	---	12.48	3.75	---	12.30	6.70
Root-mean-square error	.05	.06	.06	.08	.08	.08	.07	.07	.07
RMSEA	.07	.07	.07	.10	.09	.10	.07	.06	.06
Goodness of fit index	.94	.93	.93	.88	.90	.88	.91	.92	.92
Comparative fit index	.95	.95	.95	.87	.90	.87	.92	.94	.93
Relative fit index	.84	.85	.84	.71	.71	.70	.73	.75	.74
Power at .05 significance level	.96	.97	.97	.99	.96	.99	.75	.58	.64

Notes: RMSEA=Root-mean-square error of approximation. The degrees of freedom for $\Delta\chi^2$ between one and four-factor model is df=6; for $\Delta\chi^2$ between one and two factor model, df=3. * p < .05.

Table 4

Correlations of Core Self-Evaluations Scale (CSES) with Other Variables

	Sample 1		Sample 2		Sample 3		Sample 4	
	r	r_c	r	r_c	r	r_c	r	r_c
Self-esteem	.67	.80	.72	.88	.77	.89	.79	.92
Generalized self-efficacy	.66	.78	.65	.80	.68	.80	.77	.90
Neuroticism	-.60	-.70	-.70	-.84	-.71	-.82	-.58	-.67
Locus of control	.35	.50	.35	.46	.42	.62	.29	.41
Conscientiousness	.53	.62	.48	.57	.47	.54	.27	.31
Extraversion	.42	.52	.41	.51	.39	.44	.42	.53
Agreeableness	---	---	.12	.16	---	---	.15	.16
Openness	---	---	-.05	-.06	---	---	-.10	-.13
Job satisfaction	.41	.49	.49	.59	---	---	---	---
Job performance	.23	.26	.24	.27	---	---	---	---
Life satisfaction	---	---	.54	.66	.45	.53	.54	.64

Notes: r=uncorrected correlation. r_c=corrected correlation.

Table 5

Usefulness Analyses of CSES Compared to Core Self-Evaluations Factor and Five-Factor Model

	Sample 1		Sample 2			Sample 3	Sample 4
	Job Satisfaction	Job Performance	Job Satisfaction	Job Performance	Life Satisfaction	Life Satisfaction	Life Satisfaction
1. CSE factor (1)	.46**	.18*	.43**	.18*	.49**	.45**	.59**
2. CSES (1)	.01*	.05*	.07**	.05*	.06**	.03**	.01
1. CSES (1)	.42**	.23**	.49**	.24**	.54**	.48**	.55**
2. CSE factor (1)	.05**	.00	.00	.00	.01	.00	.05*
1. Four core traits (4)	.47**	.26**	.44**	.20	.54**	.55**	.61**
2. CSES (1)	.01	.04*	.06**	.05*	.04**	.00	.01
1. CSES (1)	.42**	.23**	.49**	.24**	.54**	.44**	.55**
2. Four core traits (4)	.06**	.07	.01	.01	.04*	.11**	.07**
1. FFM (3[1,3],5[2,4])	.49**	.18*	.46**	.29**	.56**	.49**	.60**
2. CSES (1)	.02**	.11**	.08**	.03	.04**	.04**	.02*
1. CSES (1)	.44**	.25**	.49**	.24**	.54**	.50**	.54**
2. FFM (3[1,3],5[2,4])	.07**	.04	.05*	.09	.06**	.03*	.08**

Table 5 Continues

Table 5 (Continued)

Notes: The number of variables entered into the regression on each step is provided in parentheses. CSE factor=Core self-evaluations factor. CSES=Core Self-Evaluations Scale. For Samples 1 and 3 [1,3], five-factor model (FFM) consists of neuroticism, extraversion, and conscientiousness. For Samples 2 and 4 [2,4], all five Big Five traits were used. Table entries are multiple correlations (Multiple R). Life satisfaction in Sample 3 is at Time 2.

Numbers in second stage are change in multiple correlations (ΔR). * $p < .05$. ** $p < .01$. Differences in significance between similar incremental R^2 values are due to differences in degrees of freedom or the size of the variance explained in the full model (see Cohen & Cohen [1983], p. 145).