[Practical Assessment, Research, and Evaluation](https://scholarworks.umass.edu/pare)

Volume 26 Article 6

2021

Reliability Generalization Analysis of the Core Self-Evaluations Scale

Jisoo Ock
Pusan National University

Samuel T. McAbee
Bowling Green State University

Seydahmet Ercan
Bülent Ecevit Üniversitesi

Amy Shaw
University of Macau

Frederick L. Oswald
Rice University

Practical Assessment, Research & Evaluation

A peer-reviewed electronic journal.

Volume 26 Number 6, March 2021 ISSN 1531-7714

Reliability Generalization Analysis of the Core Self-Evaluations Scale

Jisoo Ock, *Pusan National University*
Samuel T. McAbee, *Bowling Green State University*
Seydahmet Ercan, *Bülent Ecevit Üniversitesi*
Amy Shaw, *University of Macau*
Frederick L. Oswald, *Rice University*

As a multifaceted construct reflecting one's self-esteem, generalized self-efficacy, locus of control, and emotional stability, core self-evaluations (CSE) has become popular to measure in applied psychology research, especially given its conceptual importance and empirical usefulness for understanding the dispositional effects on employee attitudes and behaviors. Yet, less attention has been paid to the internal properties of its measurement, relative to its criterion-related validity evidence. Thus, we believe that it is useful and timely to report on meta-analytic evidence regarding the psychometric reliability and associated study characteristics of Core Self-Evaluations Scale (CSES; Judge et al., 2003) to inform their nature, use, and future development. Results demonstrated support for acceptable levels of coefficient alpha across measures (μ_a = .84, τ = .05). We discuss several implications for measuring CSE in a multidimensional and generalizable manner.

Introduction

In today's era of big data and machine learning, it is easy for anyone to be captivated by the fact that effective prediction is achieved using modern approaches—without understanding *how* it is being predicted. Algorithmic bias and the "black box" of machine learning algorithms are current events and issues found in the popular press that reflect this type of concern. Yet this is not a new concern. Over 70 years ago, dust-bowl empiricism was the zeitgeist in psychology, where a measure was deemed useful so long as it predicted any outcome of psychological or societal importance (e.g., Minnesota Multiphasic Personality Inventory (MMPI) items predicting clinical diagnoses). But psychometrics prevailed in psychology, where researchers focused on ensuring that their measures (a) contained content that represented their constructs of interest (e.g., motivation, personality, knowledge); (b) identified and minimized sources of systematic and random error (e.g., idiosyncratic items, subgroup differences in test content were perceived); and (c) operated in ways consistent with the construct (e.g., trait measures were stable over time, converged with similar measures, and demonstrated interpretable patterns with other constructs).

Having learned its lessons of the past, psychology now places its primacy on substance-driven research, whether it is theory-driven (Klimoski, 1993; Schmitt & Landy, 1993) or inductive in nature (McAbee et al., 2017; Spector et al., 2014), where researchers develop and make use of psychometrically sound measures that reflect their respective constructs (i.e., content is representative and reliable) and estimate latent relationships of interest (i.e., show convergent and discriminant validity at the construct level; Binning & Barrett, 1989). Errors that contaminate a measure should be eliminated whenever possible (e.g., item content is clear and understood by all; the test is designed to minimize fatigue); however, random sources of error are still unavoidable even in well-developed measures (Lord & Novick, 1968). Therefore, measurement error variance should be quantified, with sources of error variance identified whenever possible, so that the nature and effects of error are better understood. Moreover, when the random sources of error in a measure can be estimated in the form of a reliability coefficient, then observed validities that involve this measure can be psychometrically corrected for attenuation caused by these random errors of measurement (Schmidt & Hunter, 1992). More to the point, the measure's observed variance in the formula for the correlation is corrected downward, to the proportion of it that is reliable, where that proportion is the reliability coefficient. This corrected correlation will be higher in magnitude than the original but also has a correspondingly larger confidence interval (Oswald et al., 2015), similar to relationships estimated in structural equation modeling (Ledgerwood & Shrout, 2011).

Although the loss in observed validity incurred due to measurement error variance can be estimated through these straightforward psychometric corrections, deciding on which reliability coefficient is usually based on convenience (e.g., coefficient alpha) rather than the multiple options that are possible (e.g., alpha, test-retest reliability, alternate forms reliability, or some combination thereof, Le et al., 2009). Moreover, it is important to note that reliability is a property of the scores on a measure for a *particular* sample and setting, rather than a stable property of the measurement instrument itself (Thompson, 2003). Therefore, even putting sampling error variance aside, reliability estimates used to make psychometric corrections fluctuate across samples due to critical sample characteristics, such as the composition of the sample (e.g., age variation), or administration conditions (e.g., lab vs. field; Crocker & Algina, 1986; Henson, 2001). Therefore, the extent to which reliability estimates vary across studies and samples, and the characteristics that might predict some of this variance, are empirical questions worth investigating.

Reliability generalization (RG) analysis is a tool for doing just that. Just as meta-analysis is in popular and effective use to summarize effect sizes across studies (e.g., correlations, *d*-values), and what factors moderate those effects, RG is a form of meta-analysis that estimates the mean, the variance, and predictors of reliability coefficient(s) across studies. More specifically, results from RG analysis indicate whether reliability coefficients for a measure are typically high or low. When average reliability is high, that gives one some assurance (but no guarantee) that reliability will be similarly high in future studies and settings. When average reliability is low, that can call the quality of the measure into question in terms of its items converging on a stable construct (which in turn can be a cause for obtaining low validities). RG analyses also provide the variance in reliability coefficients, where reliabilities may be high in some studies but not in others. In this case, and if studies in an RG vary on characteristics that can be coded accurately (e.g., industry type, percent female, average age), then perhaps the variance in reliability coefficients can be at least partially accounted for by these study-level characteristics (Vacha-Haase, 1998).

Reliability Generalization of Core Self-Evaluations Measures

In this study, we conducted a reliability generalization (RG) analysis for measures of *core self-evaluations* (CSE), a personality trait associated with the fundamental evaluations that people hold about themselves (Judge et al., 1997). Judge and colleagues (Judge et al., 1997) introduced core self-evaluations as a unifying framework to explain dispositional effects on employee job satisfaction and job performance (Judge & Bono, 2001; Judge et al., 2002). CSE is a multidimensional construct that represents the shared variance between each of four core traits: (a) *self-esteem*, one's overall level of self-respect and self-regard; (b) *generalized self-efficacy*, one's belief or evaluation in solving problems and challenges that one faces; (c) *locus of control*, the belief that one is in control of and responsible for the events that occur in one's life; and (d) *emotional stability*, the tendency to be even-tempered, optimistic, and free of negative cognitive evaluations of the self (Judge et al., 1997; Judge et al., 1998). According to Judge and colleagues, the overlap (shared variance) across these four core traits reflects the fundamental evaluations that people hold about their own worth, confidence, and competence (Judge et al., 1997; Judge et al., 1998).

Because personality and other dispositions meaningfully affect employee attitudes and behaviors, CSE and its measurement has generated a lot of attention not only in organizations, but also in psychological research and practice. CSE has not only lived up to the promise of criterion-related validity; its most popular measure, the Core Self-Evaluation Scale (CSES; Judge et al., 2003) is a short 12-item scale and thus can be administered quickly and conveniently. Moving beyond its validity and administrative convenience, the CSES has received relatively less research attention to the reliability and internal properties of its measurement (Johnson et al., 2011; Johnson et al., 2012). This is of concern, because only if the CSES is reliable and construct-relevant are its predictions substantively meaningful (Klimoski, 1993; Schmitt & Landy, 1993). This point is related to the fundamental idea that observed measures and relationships between them are *not* the same as the intended latent constructs and relationships between them (Binning & Barrett, 1989).

We assert that a renewed emphasis on the reliability of psychological measures is critically important these days, given such recent and rising popularity of the use of complex measures that reflect multiple constructs in psychological research, such as the CSES (Johnson et al., 2011), but also situational judgment tests (SJT) and AI-based virtual reality games of the future. In this context, conceptual and measurement clarity is critical. For example, multiple constructs may overlap and be represented as a hierarchy, a bifactor model, or a network. Or as with CSES, the overlap between constructs may be emphasized in creating this short 12-item measure, thus attempting to reduce multiple constructs to one, that of core self-evaluation (with some tradeoffs in doing so, see Schmitt, 2004). Regardless of the choices one makes, accurate estimation and interpretation of the observed relationships between multidimensional constructs and criteria of interest requires developing a sound nomological net that is subject to systematic testing and data-informed revisions (Edwards, 2001; Law et al., 1998). With this context in mind, we believe that it is very timely to gather meta-analytic evidence on the psychometric reliability of the CSES measure associated study characteristics, to better understand the measure and its context for use, with implications for future research and practice.

Specifically, our reliability generalization analysis is based on 189 alpha coefficients for the Core Self-Evaluation Scale (CSES), which are examined in terms of subgroup moderators (e.g., language, delivery method, publication status), therefore usefully allowing researchers to understand how reliable their own implementation of the CSES might be. Reliability generalization is an underused yet important way to gain insights about the reliability of any psychological measure of interest, beyond any individual study taken alone. Meta-analyses tend to summarize effect sizes (e.g., validities or mean differences) regardless of the measures that were used (for an exception, see McAbee & Oswald, 2013). By contrast, RG analysis fundamentally acknowledge that measures of a construct do vary meaningfully, and therefore the nature and reliability of each measure used in research and practice should be examined more carefully.

Being a direct measure, the CSES is intended to assess core self-evaluations directly; in other words, elements of all four constructs underlying CSE—self-esteem, generalized self-efficacy, locus of control, and emotional stability—are embedded within each item, such that item cuts across multiple traits. For example, the item "I determine what will happen in my life" captures both generalized self-efficacy and locus of control (Chang et al., 2012). Again, the CSES comprises 12 items in total. Regarding the reliability of CSES scores, coefficient alpha is commonly reported as an index of internal consistency of the items. Under the assumption that a given measure is unidimensional, where constituent items largely reflect the construct of interest, then high alpha reliability indicates that variance in the scale score largely reflects variance in the construct (versus error variance; see Cortina, 1993; Cronbach, 1951). Note that alpha can under-estimate reliability when compared with omega reliability, which is estimated from the squared loadings in a factor analysis model (Cortina et al., 2020). Moreover, test-retest and alternate test forms reliability are legitimate, useful, and convergent indices to gain support that the CSE construct is stable and not confounded with any particular measure; however, these reliability indices are not as frequently reported for any measure, let alone for CSE measures (Hogan, Benjamin, & Brezinski, 2000). These should be incorporated more frequently into reliability estimation, on their own or within more integrated models and estimates of reliability (e.g., the GCES reliability estimate of Le et al., 2009).

Method

Literature Search

Our comprehensive literature search of core self-evaluations and the CSES encompassed both published and unpublished research articles. The search window started in 2003, when the CSES was introduced (Judge et al., 2003) and ended with Online First publications in June 2020. More specifically, we performed a cited reference search of the Judge et al. (2003) article using the *Web of Knowledge* database to identify primary studies. We also searched multiple databases, including *PsycINFO*, *Dissertation Abstracts*, and *Google Scholar* using relevant keywords such as *core self-evaluations* and *core self-evaluations scale*. In addition to online database searching, we manually reviewed articles in applied psychology, personality, and management journals for studies that used measures of CSE; and we reviewed and incorporated studies reported in CSE literature reviews (e.g., Johnson et al., 2008) and meta-analyses (e.g., Chang et al., 2012). Furthermore, we identified unpublished articles (e.g., unpublished thesis and dissertations, conference papers) by reviewing the reference sections of CSE literature reviews and meta-analyses and by searching the Digital Dissertation Web site and online conference programs using the keywords mentioned above. For unpublished articles, we contacted the original authors for the article or for information relevant to our analysis. If the unpublished article was later published in a refereed journal, we only included the published article into our analysis. This resulted in a very highly representative, if not fully comprehensive, set of studies within the specified time window of our search activities.

Inclusion and Exclusion Criteria

We applied inclusion and exclusion criteria to all CSES studies that we identified. Studies were included if they reported CSES alpha reliability estimates from their respective samples. Thus, we removed articles that mentioned the CSES but did not administer the scale; other articles were removed because (a) they altered the CSES for study-specific purposes (e.g., modified to measure CSE at the group level), or (b) they reported inconsistent information (e.g., inconsistent coefficient alpha). Additional articles were removed when they reported that the CSES was administered, but they did not report the alpha reliability coefficient, and this information was not provided by the original authors when requested. Ultimately, a total of 162 studies and 189 unique alpha reliability estimates with an aggregate sample size of N = 54,907 were retained for the meta-analytic procedures.

Coding Scheme and Study Characteristics

Studies were coded based on language of the measure, delivery method (online vs. paper-pencil), and publication type (published vs. unpublished). Table 1 lists the information extracted from the individual articles included in the study. Four of the current authors independently coded one quarter of the studies, and each coder evaluated one additional quarter of the studies for overlapping coding and verification. Discrepancies were minor and were readily resolved through discussion with the fifth author as an independent arbiter. For example, in a small number of instances, coders missed the information provided in the primary study or miscoded information in an open code item (see Table 1).

Meta-Analytic Procedures

As we have noted, an RG analysis is a special form of meta-analysis applied to reliability coefficients. In general, when choosing to conduct a meta-analysis, a researcher must decide whether to adopt a fixed-, random-, or mixed-effects model (Borenstein et al., 2009; Hedges & Vevea, 1998; Schmidt & Hunter, 2015; Schmidt et al., 2009). The fixed-effects meta-analysis model is the most parsimonious, by assuming a single, population effect size across all study effects, where any variability in observed effects is not substantive in nature but rather due to sampling error variance and other psychometric artifacts (Borenstein et al., 2009). The fixed-effects model might fit the data sometimes, but is highly unrealistic in its assumptions, because many substantive influences on reliability coefficients across studies are possible, such as sample demographics, temporal and group differences in perceptions of the items, the employment or research setting in which the measures are administered, and so forth. Even though such influences may not be measured, they still may be present and impart an influence on the variability in observed alpha coefficients. Thus, even though "no model is true but some are useful" (Box & Draper, 1987, p. 424), it is generally unreasonable to assume that population parameters for alpha do not vary from study to study.

Table 1. Reliability Generalization Analysis Codebook

	Code
Study descriptive information	
Full APA reference	Open
Type of publication	1 = Journal
	2 = Unpublished study/Dissertation/Master's
Scale language	Open
Delivery method	1 = Online
	2 = Paper-pencil
Sample size (*N*)	Open
CSES information	
Scale range	Open (5-point vs. 7-point scale)
Raw mean	Open
Raw *SD*	Open
Converted mean	Open (Mean converted to 5-point scale)
Converted *SD*	Open (*SD* converted to 5-point scale)
Alpha reliability	Open

Note. APA = American Psychological Association; CSES = Core Self-Evaluations Scale.

The random-effects meta-analysis model captures such variation and is generally more appropriate (Schmidt et al., 2009). Regarding the CSES, important methodological variations likely produce differential degrees of internal consistency (e.g., differences in CSES item content depending on the language of the measure). Also, empirically, our meta-analysis rejected the null hypothesis for homogeneity in alpha coefficients, $Q(188) = 1466.0$, $p < .01$, thus further supporting a random-effects model. We extended this support for random effects into a mixed-effects meta-analysis model, which models both fixed-effects and random-effects. Fixed-effects are not only reflected in the overall effect size, but also by any categorical or continuous moderator effects that are modeled (e.g., year of publication, language of the measure). After modeling the fixed-effects, any remaining variance not due to sampling error or other psychometric artifacts is estimated as random-effects variance, which is considered "true" variance due to substantive factors not identified by fixed-effect moderators (Schmidt et al., 2009).

All reliability generalization analyses of our CSES alpha reliability coefficients were conducted using Schmidt and Hunter's (2015) meta-analytic methods. All effects were weighted by the inverse of their sampling error variance, as is customary, such that all other factors held equal, alpha reliabilities based on larger sample sizes contributed more strongly to meta-analytic results than those based on smaller sample sizes.

In RG analyses, it is important to appreciate that because alpha is positive and usually more toward its ceiling of 1.0 than not (e.g., around $\alpha = .70$ or .80), the sampling error variance of alpha is generally negatively skewed. Unlike meta-analyzing correlation coefficients that are much lower (e.g., around $r = .20$ or .30), not considering this skew can bias meta-analytic results (both meta-analytic means and variances). Thus, as is customary in RG analysis, to help normalize the sampling error variance of alpha, we applied the Fisher's Z transformation (z') to alpha before estimating the mean and variance of the population reliability estimate, the associated 95% confidence interval for the mean alpha, and the 95% credibility interval that estimates the true (random-effects) variance across effect sizes. Once these statistics were computed, we back-transformed these estimates into the original alpha metric. The formula for Fisher's transformation can be written as

$$z' = (0.5)ln\frac{1+r}{1-r} \quad (1),$$

where r is the effect size to be transformed to z' (alpha coefficient in the current study). The *Z*-transformed alpha coefficients can be back-transformed to r using the following formula:

$$r = \frac{(e^{2z'} - 1)}{(e^{2z'} + 1)} \qquad (2).$$

In addition to using the formulas, it is also possible to refer to r to Z tables that are available in online statistics textbooks (e.g., https://onlinestatbook.com/2/calculators/r_to_z.html).

Moderators

Based on the aforementioned measure and study characteristics that we coded, and after using all available studies to estimate the overall reliability estimates, we conducted several subgroup analyses for categorical moderators by performing mixed-effects meta-analyses for all studies that provided a reliability estimate, in addition to information on the given moderating variable of interest.

Heterogeneity of effects

To examine the heterogeneity of our CSES alphas, we estimated the *random-effects variance* of these effects. This variance is represented by tau-squared (τ^2); however, we report the standard deviation tau (τ), because it is more interpretable (the *SD* in the metric of alpha, not in the squared metric of a variance) and it can be used to build 95% *credibility intervals* around the mean of each set of alphas of interest to indicate the 'true' range of alpha across studies. If tau and the corresponding credibility interval is sufficiently large, it suggests a large amount of heterogeneity of alpha coefficients across multiple subpopulations; whereas a small credibility interval essentially supports the assumption of the fixed effects model, that there is a single population alpha (or a small range) underlying the observed distribution of alpha coefficients (Pearlman et al., 1980).

Fixed-effect moderators can also partially account for observed heterogeneity of alphas. For example, given a categorial moderator, there can be mean differences in effects (e.g., the mean for English measures vs. mean of non-English measures) that might account for subgroup differences. Significant variance of the means, as reflected by smaller 95% confidence intervals around each mean estimate, would support the hypothesis that moderating factors are in operation. In addition, confidence intervals at each level of the moderator indicate the accuracy of the estimate of the mean effect size within each subpopulation (Whitener, 1990), keeping in mind that associated heterogeneity may still be present.

All RG analyses were conducted using an Excel spreadsheet developed by the fifth author, which is freely available at https://osf.io/gk9zr/?view_only=6fdac9798ec446f69fc5971fbdc14ee0 . However, the reader interested in RG and meta-analysis might also consider several user-friendly meta-analysis packages available for free in open-source statistical environments that can be used to conduct meta-analysis (e.g., Polanin et al., 2017, review several popular meta-analysis packages available in R).

Results

Table 2 summarizes the meta-analytic reliability estimates for the CSES measure. As would be expected in RG studies of psychometrically sound measures, the weighted population mean of the distribution of alpha coefficients for CSES scores was high (μ_α= .84; τ = .05). Generally speaking, values of alpha on the lower end of this range are often useful for research seeking to understand general relationships between variables, and in fact α = .70 has been a longstanding rule of thumb in psychological research for minimally acceptable alpha levels, despite the rule being arbitrary. Values of alpha on the higher end of this range (α = .80 to .90) would be more critical in settings where individual-level decisions were being made, such as personnel selection settings and college admissions (Lance et al., 2006; Nunnally, 1978).

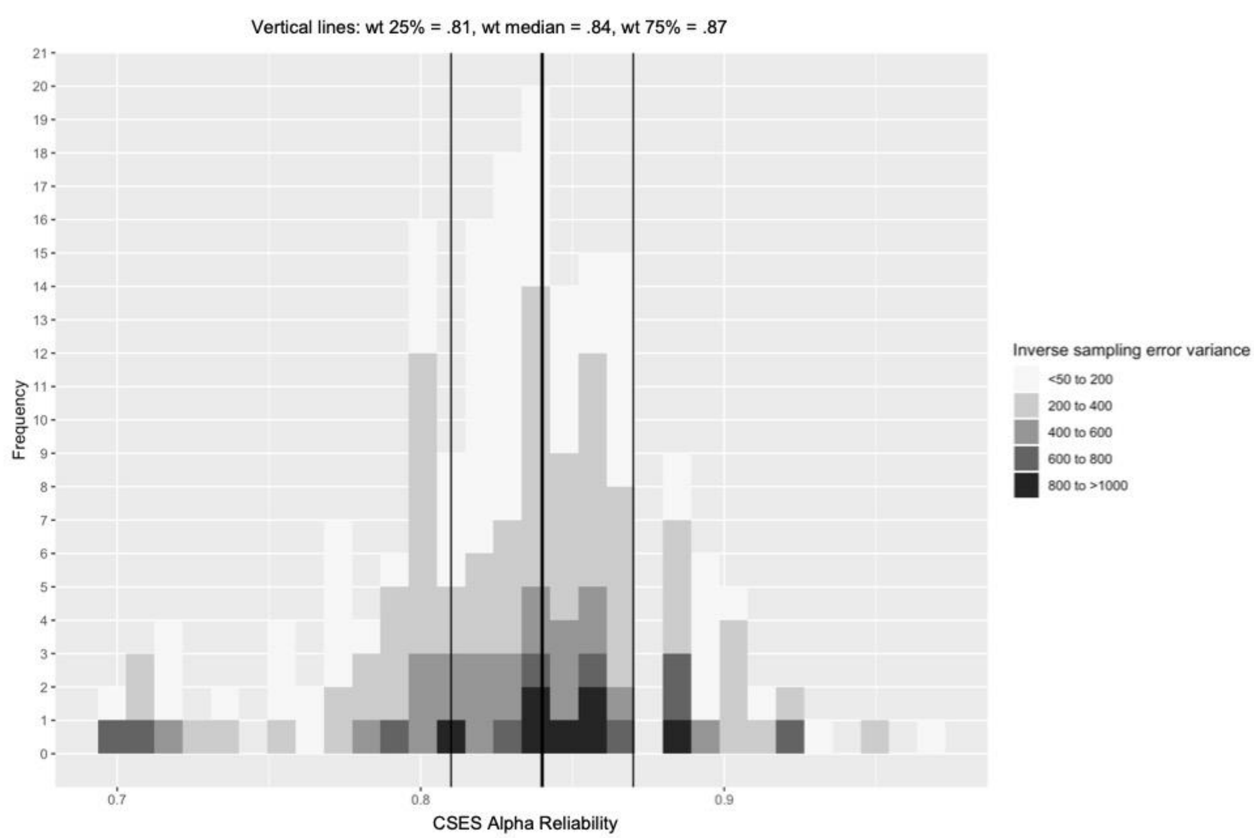
Alpha coefficients in our CSES analysis are depicted in Figure 1 as a weighted frequency distribution (Oswald & Ercan, 2013). More accurate alphas (lower sampling error variance) are darker, so that one's eyes are attracted to that part of the effect-size distribution. As the figures show, more accurate effects were generally clustered at the center of the distribution, giving some indication that stable effects were more toward the mean (less extreme), as might be expected when random effects are truly random, even if they are not perfectly normally distributed (as is assumed by 95% credibility intervals).

Table 2. CSES Measurement: Alpha Reliability Estimates

Subgroup	*N*	*k*	Mean α	*τ*	95% CI	95% CR
Overall	54,907	189	.84	.05	[.83, .85]	[.72, .91]
Language						
English	29,539	111	.85	.04	[.84, .86]	[.76, .91]
Non-English†	23,856	70	.83	.05	[.82, .84]	[.70, .91]
German	6,970	17	.85	.03	[.84, .87]	[.78, .90]
Chinese	6,003	22	.80	.07	[.77, .82]	[.64, .89]
Spanish	3,396	7	.82	.07	[.77, .87]	[.64, .91]
Korean	2,118	7	.83	.04	[.79, .85]	[.73, .89]
Delivery Method						
Online	22,402	79	.85	.04	[.84, .86]	[.75, .91]
Paper-pencil	14,471	58	.83	.06	[.81, .85]	[.67, .92]
Publication						
Published	45,743	150	.84	.05	[.83, .85]	[.72, .91]
Unpublished	9,164	36	.83	.05	[.82, .85]	[.71, .90]

Note. CSES = Core Self-Evaluations-Scale; *k* = number of effects; CI = confidence interval; CR = credibility interval. † Non-English version of CSES included the four subcategories, in addition to Romanian, Dutch, Greek, Italian, French, Finnish, Norwegian, and Persian.

Figure 1. Weighted frequency distribution of alpha coefficients for CSES measures. Darker colors represent alphas that are more precise (lower sampling error variance). Vertical lines represent weighted quartiles (25%, 50%, and 75%iles).



Regarding the RG subgroup analyses, the CSES showed similar average levels of reliability between English and non-English versions (μ_α = .85 and .83, respectively). However, note that the credibility intervals suggest that mean reliability is accompanied by some heterogeneity in the English version of the CSES, with 95% CR [.76, .91], as well as the non-English CSES, with 95% CR [.70, .91]. Table 1 shows that a statistical distinction could be made between the reliability of CSES administered in English (μ_α = .85) and German (μ_α = .85) vs. Chinese (μ_α = .80); however, even with these distinctions, mean reliability estimates across languages were still high, generally exceeding .80. Mean reliability estimates were also similarly high across delivery method (online vs. paper-pencil; μ_α = .85 and .83 respectively) and publication status (published vs. unpublished; μ_α= .84 and .83, respectively) with overlapping confidence intervals across moderating conditions, indicating that CSES scores tend to be highly reliable in terms of internal consistency, regardless of delivery method or publication status.

Without conducting this reliability generalization analysis, we would not have known that the CSES has demonstrated high levels of reliability across a wide range of studies, with minimal differences between subgroups. Also note that although reliability was generally high, reliabilities still showed meaningful levels of heterogeneity (e.g., overall 95% credibility interval was .72 to .91). All together, these findings support an important recommendation that applies to any study involving psychological measurement: Researchers should consider the distribution of reliability coefficients obtained in the past, because like a Bayesian prior, it suggests the typical reliabilities that might be obtained in the future and thus inform whether a measure should be chosen and used. Once the measure is chosen, the reliability of measures based the local sample and setting should still be calculated, which can also be used to update the past RG results.

Discussion

In light of the growing popularity of core self-evaluations and the CSES measure in psychological research, our results provide evidence that the CSES tends to be highly reliable across samples, especially when the goal is to use CSES to examine how core self-evaluations relates to other variables at the overall level (correlations, structural equation models). If the goal instead was to use CSES to make reliable decisions about individuals (e.g., personnel selection, promotion), then the standard error of measurement becomes relevant, and levels of reliability need to be as high as possible to distinguish scores from one another. Although heterogeneity of alpha was relatively small (τ = .05), it can make a difference when it comes to the aforementioned purposes of a measure.

In terms of the CSES across different languages, the mean differences and associated heterogeneity in reliability estimates across CSES languages may be due to the difference in the nature of the translated measures. For example, most of the primary studies with a German sample retained in the current RG analysis used the German versions of the CSES that are well established and frequently used (e.g., Strumpp et al., 2010). Conversely, studies using the Chinese or Spanish versions of CSES often relied on independently translated measures. This increases the likelihood of differences in the interpretation of measurement items across different language sub-populations. Although we found that non-English CSES measures were generally reliable, we nonetheless encourage future CSES researchers with non-English-speaking samples to use (or create) translated CSES that have demonstrated satisfactory psychometric properties and construct validity evidence, instead of creating or relying on independently translated measures that lack sufficient evidence for use. We contend that this practice will help produce more consistently reliable scores in non-English versions of CSES; and if the same measure is used across studies, this allows for better cross-study comparisons of results. For example, we note that recent versions of the CSES have been published in Chinese (Sun & Jiang, 2017) and Spanish (Beléndez et al., 2018). Considering the increasing popularity of CSE research around the world, research efforts to develop valid translations of CSES should continue.

Limitations and Future Directions

The reliability generalization results for the CSES, while useful for research and practice, should also be considered in the context of two limitations that are true for any psychological measure of this nature. First, although alpha is the most commonly reported reliability coefficient, we are aware that its assumptions are often violated to some extent, which can negatively or positively bias alpha as an estimate of true reliability (Cortina, 1993). Specifically, the assumptions for alpha include: (1) tau-equivalence (i.e., equivalent indicator factor loadings); (2) independent error variances (i.e., uncorrelated indicator residuals); and (3) unidimensionality (i.e., one-factor model appropriately represents the data. Second, for any self-report measure, it is psychometrically challenging to disentangle the shared variance due to actual construct overlap (which is desirable) versus the shared variance due to halo effects, implicit theories of oneself, and other measurement artifacts (which is undesirable; Podsakoff et al., 2003).

Our RG results should also be considered alongside some issues that are unique to the CSES. First, it is worth explicitly noting that the CSES reflects four CSE traits (again: self-esteem, generalized self-efficacy, locus of control, and emotional stability), which at first glance suggests that CSES is multidimensional, violating the unidimensionality assumption behind alpha. However, CSE is based on a high level of commonality (overlap) between the four traits, and the short 12-item CSES intends to capture this commonality in a single CSE dimension. Another way to measure CSE in a more refined manner, if testing time permits, is by measuring each dimension reliably and modeling CSE as a hierarchical construct. Composite reliability can then be calculated for the general CSE factor, with the four CSE traits as lower-order factors (and their own reliability coefficients; McDonald, 1999; Raykov, 1997). In addition to modeling multidimensionality, estimating the factor loadings relaxes the tau-equivalence assumption of alpha.

Second, whether a factor analysis is applied to the four factors of the CSE, or to the short CSES measure, the factor analytic approach suggests that reliability generalization across studies can be usefully refined. A collection of similar factor model estimates can be synthesized using meta-analytic structural equation modeling (MASEM) to arrive at the population composite reliability and the variability in composite reliability (see Scherer & Teo, 2020, for a tutorial on MASEM approach to meta-analyzing reliability coefficients). This might provide additional item-level information about how CSE measures vary across studies and, practically speaking, how MASEM reliability estimates differ from alpha reliability estimates (see Peterson & Kim, 2012 for an example). This may not only be a useful future direction not just for assessing reliability estimation for various measures of CSE; it may also be helpful for higher-order multidimensional constructs in general.

Our third point is more substantive, namely that extensions of the current study could examine whether the four CSE traits satisfy the theoretical and empirical inclusion criteria as indicators of the higher-order CSE construct. Specifically, the appropriateness of locus of control as a reflective indicator of CSE has been called into question in several papers (e.g., Chen, 2012; Johnson et al., 2008; Johnson et al., 2015). Theoretically, locus of control does not cleanly fit the criterion of being self-evaluative, because compared with other CSE traits that are more influenced by how one fundamentally evaluates the self, locus of control is considered to be more influenced by how one evaluates his/her environment (Johnson et al., 2015). Moreover, although high levels of self-esteem, generalized self-efficacy, and emotional stability generally seem to enhance feelings of self-worth, the relationship between high level of locus of control (i.e., internal locus of control) and self-worth might vary (Johnson et al., 2016). For example, when experiencing failure, those with high internal locus of control may feel more negative feelings of self-worth, in part due to the perception (rightly or wrongly) that one had the control to do something differently and avoid failure.

Empirically, studies corroborate these ideas, in that locus of control is relatively weakly correlated with self-esteem, generalized self-efficacy, and emotional stability (Johnson et al., 2016; Judge et al., 2002), and factor loadings emanating from the CSE factor onto locus of control tends to be weaker than the other CSE indicator traits (Dormann et al., 2006; Erez & Judge, 2001; Heller et al., 2002). Empirical evidence regarding the lack of interchangeability of locus of control with the other CSE traits is more evident when researchers attempt to control for the effect of common method variance (Johnson et al., 2016; Johnson et al., 2011). Thus, as previous researchers have mentioned (Chen, 2012; Johnson et al., 2011), future researchers should continue to refine the theoretical and empirical inclusion criteria for multidimensional constructs (CSE and otherwise) and their representation (hierarchical and otherwise).

Fourth and finally, some conditions of our reliability generalization analysis had only a small number of effects, drawing attention to the need for more studies in certain moderator conditions where we would like to understand CSES further (e.g., Spanish and Korean versions of the CSES). Additionally, because some studies used different versions of CSES (e.g., CSES translated to different languages), alpha coefficients may have been influenced by measurement-specific factors that render them less comparable. To some extent, we were able to model this heterogeneity meta-analytically, by way of estimating moderator effects and random-effects variance; but specific differences in CSES across studies (e.g., alterations or deletions of specific items) should be considered more closely and substantively as well. And as always, new moderators can be investigated (e.g., effects by gender, age, and type of industry).

Conclusion

Psychologists are constantly measuring people's thoughts and behaviors, but this is only useful if the measures are highly construct-relevant and the data provided from those measures are reliable. Reliable measures are important for research purposes, and reliability is especially important in applied contexts where test scores may be used to make high-stakes decisions that have important consequences for both the individual (e.g., selection, promotion) and the organization. We found that the alpha coefficient of CSES scores, on average, is generally high across studies. Although we did find some differences in the mean and variance of the effects with respect to the language in which the CSES was measured, they were not high enough to cause practical concerns when deciding on whether to use the CSES. These reliability generalization results are heartening overall, where the alpha reliability of CSES scores have generally been high and can therefore be generally expected to be high in future studies. If a future study shows very low reliability for the CSES, this would be unexpected, and one might follow up to determine whether the CSES was scored correctly, or whether there were sample characteristics that deviated meaningfully from the collective of samples within the current RG analysis.

In closing, however, we wish to emphasize that very important questions regarding the nature, reliability, and validation of measures of multidimensional constructs remain. Regarding CSE measures, for example, research can continue to investigate whether the short CSES measure captures the general CSE construct in a manner that represents the lower-order constructs as intended; and it might continue to be usefully pitted against measures of its lower-order constructs in terms of criterion-related validity and broader patterns of convergent and discriminant validity. The current reliability generalization analysis, and these future research directions inspired by it, should serve as a model for further examining other higher-order multidimensional constructs beyond CSE, such as psychological capital; and in turn, lessons learned about multidimensional measurement in the domain of cognitive ability can likely be applied in multiple domains of non-cognitive measurement where multidimensionality is a reality and a challenge.

References

Studies included in the meta-analysis are indicated by an asterisk.

*Ahmetoglu, G., Leutner, F., & Chamorro-Premuzic, T. (2011). EQ-nomics: Understanding the relationship between individual differences in trait emotional intelligence and entrepreneurship. *Personality and Individual Differences*, *51*, 1028-1033. http://dx.doi.org/10.1016/j.paid.2011.08.016

*Ahn, J., Lee, S., & Yun, S. (2018). Leaders' core self-evaluation, ethical leadership, and employees' job performance: The moderating role of employees' exchange ideology. *Journal of Business Ethics, 148,* 457-470. http://dx.doi.org/10.1007/s10551-016-3030-0

*Avey, J. B., Luthans, F., & Youssef, C. M. (2010). The additive value of positive psychological capital in predicting work attitudes and behaviors. *Journal of Management*, *36*, 430-452. http://dx.doi.org/10.1177/0149206308329961

*Beal, D. J., Trougakos, J. P., Weiss, H. M., & Green, S. G. (2006). Episodic processes in emotional labor: Perceptions of affective delivery and regulation strategies. *Journal of Applied Psychology*, *91*, 1053-1065. http://dx.doi.org/10.1037/0021-9010.91.5.1053

* Beléndez, M., Gómez, A., López, S., & Topa, G. (2018). Psychometric properties of the Spanish version of the Core Self-Evaluations Scale (CSES-SP). *Personality and Individual Differences, 122,* 195-197. http://dx.doi.org/10.1016/j.paid.2017.10.034

Binning, J. F., & Barrett, G. V. (1989). Validity of personnel decisions: A conceptual analysis of the inferential and evidential bases. *Journal of Applied Psychology, 74,* 478-494. http://dx.doi.org/10.1037/0021-9010.74.3.478

*Blau, G. (2007). Partially testing a process model for understanding victim responses to an anticipated worksite closure. *Journal of Vocational Behavior*, *71*, 401-428. http://dx.doi.org/10.1016/j.jvb.2007.08.005

*Blume, B. D., Baldwin, T. T., & Rubin, R. S. (2005). *Forced raking: Who is attracted to it? A study of performance management system preference*. Paper presented at the 65th Academy of Management Annual Meeting, Honolulu, HI.

*Boekhorst, J. A. (2015). *Human resource management practices, work intensity, and workplace deviance: Exploring the moderating role of core self-evaluations* (Unpublished doctoral dissertation). York University, Toronto, ON.

*Booth, J. E., Shantz, A., Glomb, T. M., Duffy, M. K., & Stillwell, E. E. (2020). Bad bosses and self-verification: The moderating role of core self-evaluations with trust in workplace management. *Human Resource Management, 59,* 135-152. http://dx.doi.org/10.1002/hrm.21982

Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2009). *Introduction to meta-analysis*. Hoboken, NJ: Wiley.

*Bowling, N. A. (2005). *Dispositional approach to satisfaction with social facets of work: The role of dispositional cynicism* (Unpublished doctoral dissertation). Central Michigan University, Mount Pleasant, MI.

*Bowling, N. A., Wang, Q., & Li, H. Y. (2012). The moderating effect of core self-evaluations on the relationships between job attitudes and organizational citizenship behavior. *Applied Psychology: An International Review*, *61*, 97-113. http://dx.doi.org/10.1111/j.1464-0597.2011.00458.x

Box, G. E. P., & Draper, N. R. (1987). *Empirical model-building and response surfaces*. New York: Wiley.

*Brown, D. J., Ferris, D. L., Heller, D., & Keeping, L. M. (2007). Antecedents and consequences of the frequency of upward and downward social comparisons at work. *Organizational Behavior and Human Decision Processes*, *102*, 59-75. http://dx.doi.org/10.1016/j.obhdp.2006.10.003

*Brunborg, G. S. (2008). Core self-evaluations. *European Psychologist*, *13*, 96-102. http://dx.doi.org/10.1027/1016-9040.13.2.96

*Cable, D. M., & Kay, V. S. (2012). Striving for self-verification during organizational entry. *Academy of Management*, *55*, 360-380. http://dx.doi.org/10.5465/amj.2010.0397

*Campana, K. L., & Bono, J. E. (2007). *Negative work events, mood, and motivation: The role of core self-evaluations.* Paper presented at the 22nd Annual Meeting of the Society for Industrial and Organizational Psychology, New York, NY.

*Cassar, G., & Friedman, H. (2009). Does self-efficacy affect entrepreneurial investment? *Strategic Entrepreneurship Journal*, *3*, 241-260. http://dx.doi.org/10.1002/sej.73

*Chamorro-Premuzic, T., Ahmetoglu, G., & Furnham, A. (2008). Little more than personality: Dispositional determinants of test anxiety (the Big Five, core self-evaluations, and self-assessed intelligence). *Learning and Individual Differences*, *18*, 258- 263. http://dx.doi.org/10.1016/j.lindif.2007.09.002

Chang, C. –C., Ferris, D. L., Johnson, R. E., Rosen, C. C., & Tan, J. A. (2012). Core self-evaluations: A review and evaluation of the literature. *Journal of Management*, *38*, 81-128. http://dx.doi.org/10.1177/0149206311419661

*Chase, W. S., II. (2009). *Trait mindfulness, overall adjustment, and physical well-being perceptions of college students* (Unpublished doctoral dissertation). Marywood University, Scranton, PA.

Chen, G. (2012). Evaluating the core: Critical assessment of core self-evaluations theory. *Journal of Organizational Behavior, 33,* 153-160. http://dx.doi.org/10.1002/job.761

*Chi, N. W., Yang, J., & Lin, C. Y. (2018). Service workers' chain reactions to daily customer mistreatment: Behavioral linkages, mechanisms, and boundary conditions. *Journal of Occupational*

Health Psychology, 23, 58-70. http://dx.doi.org/10.1037/ocp0000050

*Chiang, Y. H., Hsu, C. C., & Hung, K. P. (2014). Core self-evaluation and workplace creativity. *Journal of Business Research, 67,* 1405-1413. http://dx.doi.org/10.1016/j.jbusres.2013.08.012

*Chng, D. H. M., Rodgers, M. S., Shih, E., & Song, X. B. (2012). When does incentive compensation motivate managerial behaviors? An experimental investigation of the fit between incentive compensation, executive core self-evaluation, and firm performance. *Strategic Management, 33,* 1343-1362. http://dx.doi.org/10.1002/smj.1981

*Coder, L. (2007). *Understanding voluntary turnover: Comparing factors of employer change and career change in a sample of professional workers* (Unpublished doctoral dissertation). University of Kansas, Lawrence, KS.

*Cornell, N. K. (2012). *Organizational culture incongruence and employee outcomes: Examining the effects of organizational value importance, self-esteem, and stress* (Unpublished doctoral dissertation). University of Calgary, Calgary, AB.

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, *78*, 98-104. http://dx.doi.org/10.1037/0021-9010.78.1.98

Cortina, J. M., Sheng, Z., Keener, S. K., Keeler, K. R., Grubb, L. K., Schmitt, N., Tonidandel, S., Summerville, K. M., Heggestad, E. D., & Banks, G. C. (2020). From alpha to omega and beyond! A look at the past, present, and (possible) future of psychometric soundness in the Journal of Applied Psychology. *Journal of Applied Psychology, 105,* 1351-1381. http://dx.doi.org/10.1037/apl0000815

Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory.* New York, NY: Hold, Reinhard & Winston.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*, 297-334. http://dx.doi.org/10.1007/BF02310555

*Dasborough, M., Harvey, P., & Stillman, P. (2007). *Follower undermining of leaders in the workplace.* Paper presented at the 22nd Annual Meeting of the Society for Industrial and Organizational Psychology, New York, NY.

*Debusscher, J., Hofmans, J., & De Fruyt, F. (2017). Core self-evaluations as a moderator of the relationship between task complexity, job resources, and performance. *European Journal of Work and Organizational Psychology, 26,* 411-420. http://dx.doi.org/10.1080/1359432X.2016.1277706

*Dietl, E., & Meurs, J. A. (2019). Implicit core self-evaluations and work outcomes: Validating an indirect measure. *Journal of Occupational and Organizational Psychology.* http://dx.doi.org/10.1111/joop.12244

*Dietl, E., Rule, N., & Blickle, G. (2018). Core self-evaluations mediate the association between leaders' facial appearance and their professional success: Adults' and children's perspective. *Leadership Quarterly, 29,* 476-488. http://dx.doi.org/10.1016/j.leaqua.2018.01.002

*Di Fabio, A., & Palazzeschi, L. (2012). Incremental variance of the core self-evaluation construct compared to fluid intelligence and personality traits in aspects of decision-making. *Personality and Individual Differences*, *53*, 196-201. http://dx.doi.org/10.1016/j.paid.2012.03.012

Dormann, C., Fay, D., Zapf, D., & Frese, M. (2006). A state-trait analysis of job satisfaction: On the effect of core self-evaluations. *Applied Psychology: An International Review*, *55*, 27-51. http://dx.doi.org/10.1111/j.1464-0597.2006.00227.x

*Duffy, R. D., Allan, B. A., & Bott, E. M. (2012). Calling and life satisfaction among undergraduate students: Investigating mediators and moderators. *Journal of Happiness Studies*, *13*, 469-479. http://dx.doi.org/10.1007/s10902-011-9274-6

*Duffy, R. D., Diemer, M. A., Perry, J. C., Laurenzi, C., & Torrey, C. L. (2012). The construction and initial validation of the Work Volition Scale. *Journal of Vocational Behavior*, *80*, 400-411. http://dx.doi.org/10.1016/j.jvb.2011.04.002

Edwards, J. R. (2001). Multidimensional constructs in organizational behavior research: An integrative analytical framework. *Organizational Research Methods*, *4*, 144-192. http://dx.doi.org/10.1177/109442810142004

*Ellis, L. (2015). *Tomorrow's leading ladies: How core self-evaluations, leadership perceptions, and social support influence young women's leadership aspirations* (Unpublished doctoral dissertation). Clemson University, Clemson, SC.

Erez, A., & Judge, T. A. (2001). Relationship of core self-evaluations to goal setting, motivation, and performance. *Journal of Applied Psychology*, *86*, 1270-1279. http://dx.doi.org/10.1037/0021-9010.86.6.1270

*Eschleman, K. J., & Burns, G. (2012). The utility of general and school-specific personality traits and an examination of their relationship over time. *Learning and Individual Differences*, *22*, 537-543. http://dx.doi.org/10.1016/j.lindif.2012.03.007

*Extremera, N., & Rey, L. (2018). Core self-evaluations are associated with judgments of satisfaction with life via positive but not negative affect. *Personality and Individual Differences, 130,* 112-116. http://dx.doi.org/10.1016/j.paid.2018.03.054

*Fang, R., McAllister, D. J., & Duffy, M. K. (2017). Down but not out: Newcomers can compensate for low vertical access with strong horizontal ties and favorable core self-evaluations. *Personnel Psychology, 70,* 517-555. http://dx.doi.org/10.1111/peps.12177

*Ferris, D. L., Rosen, C. R., Johnson, R. E., Brown, D. J., Risavy, S. D., & Heller, D. (2011). Approach or avoidance (or both?): Integrating core self-evaluations within an approach/avoidance framework. *Personnel Psychology*, *64*, 137-161. http://dx.doi.org/10.1111/j.1744-6570.2010.01204.x

*Feys, M., Anseel, F., & Willie, B. (2011). Improving feedback reports: The role of procedural information and information specificity. *Academy of Management Learning & Education*, *10*, 661-681. http://dx.doi.org/10.5465/amle.2010.0054

*Flynn, C. B., Smither, J. W., & Walker, A. G. (2016). Exploring the relationship between leaders' core self-evaluations and subordinates' perceptions of servant leadership. *Journal of Leadership & Organizational Studies, 23,* 260-271. http://dx.doi.org/10.1177/1548051815621257

*Galvin, B. M., Waldman, D. A., & Balthazard, P. (2010). Visionary communication qualities as mediators of the relationship between narcissism and attributions of leader charisma. *Personnel Psychology*, *63*, 509-537. http://dx.doi.org/10.1111/j.1744-6570.2010.01179.x

*Garden, R., Hu, X., Zhan, Y., & Yao, X. (2018). Popularity procurement and pay off: Antecedents and consequences of popularity in the workplace. *Journal of Business and Psychology, 33,* 297-310. http://dx.doi.org/10.1007/s10869-017-9494-9

*Gardner, D. G., & Pierce, J. L. (2010). The Core Self-Evaluation Scale: Further construct validation evidence. *Educational and Psychological Measurement*, *70*, 291-304. http://dx.doi.org/10.1177/0013164409344505

*Grant, A. M., & Sonnentag, S. (2010). Doing good buffers against feeling bad: Prosocial impact compensates for negative task and self-evaluations. *Organizational Behavior and Human Decision Processes*, *111*, 13-22. http://dx.doi.org/10.1016/j.obhdp.2009.07.003

*Grant, A. M., & Wrzesniewski, A. (2010). I won't let you down… or will I? Core self-evaluations, other-orientation, anticipated guilt and gratitude, and job performance. *Journal of Applied Psychology*, *95*, 108-121. http://dx.doi.org/10.1037/a0017974

*Greaves, C. E., Parker, S. L., Zacher, H., & Jimmieson, N. L. (2017). Working mothers' emotional exhaustion from work and care: The role of core self-evaluations, mental health, and control. *Work & Stress, 31,* 164-181. http://dx.doi.org/10.1080/02678373.2017.1303760

*Greenbaum, R. L., Mawritz, M. B., & Eissa, G. (2012). Bottom-line mentality as an antecedent of social undermining and the moderating roles of core self-evaluations and conscientiousness. *Journal of Applied Psychology*, *97*, 343-359. http://dx.doi.org/10.1037/a0025217

*Greguras, G. J., & Diefendorff, J. M. (2008). *Proactive personality, core self-evaluations, and self-determination theory.* Paper presented at the 23rd Annual Meeting of the Society of Industrial and Organizational Psychology, San Francisco, CA.

*Grosser, T. J., Obstfeld, D., Labianca, G., & Borgatti, S. (in press) Measuring mediation and separation brokerage orientations: A further step toward studying the social network brokerage process. *Academy of Management Discoveries.* http://dx.doi.org/10.5465/amd.2017.0110

*Gu, H., Wen, Z., & Fan, X. (2015). The impact of wording effect on reliability and validity of the core self-evaluations scale (CSES): A bi-factor perspective. *Personality and Individual Differences, 83,* 142-147. http://dx.doi.org/10.1016/j.paid.2015.04.006

*Guan, Y., Dai, X., Gong, Q., Deng, Y., Hou, Y., Dong, Z., … Lai, X. (2017). Understanding the trait basis of career adaptability: A two-wave mediation analysis among Chinese university students. *Journal of Vocational Behavior, 101,* 32-42. http://dx.doi.org/10.1016/j.jvb.2017.04.004

*Guo, Y., Zhu, Y., & Zhang, L. (in press). Effect of core self-evaluations on research performance among Chinese university teachers. *Current Psychology*. http://dx.doi.org/10.1007/s12144-017-9685-6

*Hannah, S. T., Avolio, B. J., Walumbwa, F. O., & Chan, A. (2012). Leader self and means efficacy: A multi-component approach. *Organizational Behavior and Human Decision Processes*, *118*, 143-161. http://dx.doi.org/10.1016/j.obhdp.2012.03.007

*Harris, K. J., Harvey, P., & Kacmar, K. M. (2009). Do social stressors impact everyone equally? An examination of the moderating impact of core self-evaluations. *Journal of Business and Psychology*, *24*, 153-164. http://dx.doi.org/10.1007/s10869-009-9096-2

*Haskins, A., Roebke, M. A., Woolley, M. R., Nordström, T., & Burns, G. N. (2018). *Role conflict and core self-evaluations: Effects of stress appraisals on well-being.* Poster presented at the 33rd Annual Conference of the Society for Industrial and Organizational Psychology, Chicago, IL.

*Haynie, J. J. (2012). Core self-evaluations and team performance: The role of team-member exchange. *Small Group Research*, *43*, 315-329. http://dx.doi.org/10.1177/1046496411428357

Hedges, L. V., & Vevea, J. L. (1998). Fixed- and random-effects models in meta-analysis. *Psychological Methods*, *3*, 486-504. http://dx.doi.org/10.1037/1082-989X.3.4.486

*Heilmann, T., & Jonas, K. (2010). Validation of a German-language Core Self-Evaluations Scale. *Social Behavior and Personality*, *38*, 209-226. http://dx.doi.org/10.2224/sbp.2010.38.2.209

Heller, D., Judge, T. A., & Watson, D. (2002). The confounding role of personality and trait affectivity in the relationship between job and life satisfaction. *Journal of Organizational Behavior*, *23*, 815-835. http://dx.doi.org/10.1002/job.168

Henson, R. K. (2001). Understanding internal consistency reliability estimates: A conceptual primer on coefficient alpha. *Measurement and Evaluation in Counseling and Development, 34,* 177-189. http://dx.doi.org/10.1080/07481756.2002.12069034

*Henson, J. A., & Beehr, T. (2018). Subordinates' core self-evaluations and performance predict leader-rated LMX. *Leadership & Organization Development Journal, 39,* 150-168. http://dx.doi.org/10.1108/LODJ-06-2016-0162

*Hentrich, S., Zimber, A., Sosnowsky-Waschek, N., Gregersen, S., & Petermann, F. (2017). The role of core self-evaluations in explaining depression and work engagement among managers. *Current Psychology, 36,* 516-529. http://dx.doi.org/10.1007/s12144-016-9439-x

*Herrmann, A., Hirschi, A., & Baruch, Y. (2015). The protean career orientation as predictor of career outcomes: Evaluation of incremental validity and mediation effects. *Journal of Vocational Behavior, 88,* 205-214. http://dx.doi.org/10.1016/j.jvb.2015.03.008

*Hirschi, A. (2011). Callings in career: A typological approach to essential and optional components. *Journal of Vocational Behavior*, *79*, 60-73. http://dx.doi.org/10.1016/j.jvb.2010.11.002

*Hirschi, A., Herrmann, A., & Keller, A. C. (2015). Career adaptability and adapting: A conceptual and empirical investigation. *Journal of Vocational Behavior, 87,* 1-10. http://dx.doi.org/10.1016/j.jvb.2014.11.008

Hogan, T. P., Benjamin, A., & Brezinski, K. L. (2000). Reliability methods: A note on the frequency of use of various types. *Educational and Psychological Measurement, 60,* 523-531. http://dx.doi.org/10.1177/00131640021970691

*Holtschlag, C., Reiche, B. S., & Masuda, A. D. (2018). How and when do core self-evaluations predict career satisfaction? The roles of positive goal emotions and occupational embeddedness. *Applied Psychology: An International Review, 68,* 126-149. http://dx.doi.org/10.1111/apps.12153

*Howard, M. C. (2017). The empirical distinction of core self-evaluations and psychological capital and the identification of negative core self-evaluations and negative psychological capital. *Personality and Individual Differences, 114,* 108-118. http://dx.doi.org/10.1016/j.paid.2017.03.061

*Hsieh, H. H., & Huang, J. T. (2017). Core self-evaluations and job and life satisfaction: The mediating and moderated mediating role of job insecurity. *Journal of Psychology: Interdisciplinary and Applied, 151,* 282-298. http://dx.doi.org/10.1080/00223980.2016.1270888

*Hu, J., Wang, Z., Liden, R. C., & Sun, J. (2012). The influence of leader core self-evaluation on follower reports of transformational leadership. *Leadership Quarterly*, *23*, 860-868. http://dx.doi.org/10.1016/j.leaqua.2012.05.004

*Huang, J. L., Liu, M., & Bowling, N. A. (2015). Insufficient effort responding: Examining an insidious confound in survey data. *Journal of Applied Psychology, 100,* 828-845. http://dx.doi.org/10.1037/a0038510

*Iliescu, D., Ispas, D., Sulea, C., & Ilie, A. (2015). Vocational fit and counterproductive work behaviors: A self-regulation perspective. *Journal of Applied Psychology, 100,* 21-39. http://dx.doi.org/10.1037/a0036652

*Imus, A. L. (2006). *Preferential selection in the academic domain: A stigma-based model of antecedents and outcomes* (Unpublished master's thesis). Michigan State University, East Lansing, MI.

*Ispas, D., & Ilies, A. (2008). *Negative emotions, core self-evaluations and counterproductive work behaviors*. Paper presented at the 23rd Annual Meeting of the Society for Industrial and Organizational Psychology, San Francisco, CA.

*Jiang, Z. (2015). Core self-evaluation and career decision self-efficacy: A mediation model of value orientations. *Personality and Individual Differences, 86,* 450-454. http://dx.doi.org/10.1016/j.paid.2015.07.012

*Jiang, Z., Wang, Z., Jing, X., Wallace, R., Jiang, X., & Kim, D. (2017). Core self-evaluation: Linking career social support to life satisfaction. *Personality and Individual Differences, 112,* 128-135. http://dx.doi.org/10.1016/j.paid.2017.02.070

Johnson, R. E., Rosen, C. C., & Chang, C. –H. (2011). To aggregate or not to aggregate: Steps for developing and validating higher-order multidimensional constructs. *Journal of Business and Psychology*, *26*, 241-248. http://dx.doi.org/10.1007/s10869-011-9238

Johnson, R. E., Rosen, C. C., Chang, C. –H., Djurdjevic, E., & Taing, M. U. (2012). Recommendations for improving the construct clarity of higher-order multidimensional constructs. *Human Resource Management Review*, *22*, 62-72. http://dx.doi.org/10.1016/j.hrmr.2011.11.006

Johnson, R. E., Rosen, C. C., Chang, C. –H., & Lin, S. –H. (2015). Getting to the core of locus of control: Is it an evaluation of the self or the environment. *Journal of Applied Psychology, 100,* 1568-1578. http://dx.doi.org/10.1037/apl0000011

Johnson, R. E., Rosen, C. C., Chang, C. –H., & Lin, S. –H. (2016). Assessing the status of locus of control as an indicator of core self-evaluations. *Personality and Individual Differences, 90,* 155-162. http://dx.doi.org/10.1016/j.paid.2015.11.002

Johnson, R. E., Rosen, C. C., & Djurdjevic, E. (2011). Assessing the impact of common method variance on higher order multidimensional constructs. *Journal of Applied Psychology*, *96*, 744-761. http://dx.doi.org/10.1037/a0021504

Johnson, R. E., Rosen, C. C., & Levy, P. E. (2008). Getting to the core of core self-evaluation: A review and recommendations. *Journal of Organizational Behavior*, *29*, 391-413. http://dx.doi.org/10.1002/job.514

*Joo, B. K. B., Jeung, C. W., & Yoon, H. J. (2010). Investigating the influences of core self-evaluations, job autonomy, and intrinsic motivation on in-role job performance. *Human Resource Development Quarterly*, *21*, 353-371. http://dx.doi.org/10.1002/hrdq.20053

*Joo, B. K., & Jo, S. J. (2017). The effects of perceived authentic leadership and core self-evaluations on organizational citizenship behavior. *Leadership & Organization Development Journal, 38,* 463-481. http://dx.doi.org/10.1108/LODJ-11-2015-0254

*Joo, B. K., Yoon, H. J., & Jeung, C. W. (2012). The effects of core self-evaluations and transformational leadership. *Leadership & Organization Development Journal, 33,* 564-582. http://dx.doi.org/10.1108/13620431211241090

Judge, T. A., & Bono, J. E. (2001). 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. http://dx.doi.org/10.1037/0021-9010.86.1.80

*Judge, T. A., Erez, A., Bono, J. E., & Thoresen, C. J. (2003). The Core Self-Evaluations Scale: Development of a measure. *Personnel Psychology*, *56*, 303-331. http://dx.doi.org/10.1111/j.1744-6570.2003.tb00152.x

Judge, T. A., Heller, D., & Mount, M. K. (2002). Five-factor model of personality and job satisfaction: A meta-analysis. *Journal of Applied Psychology, 87,* 530-541. http://dx.doi.org/10.1037/0021.9010.87.3.530

*Judge, T. A., Hurst, C., & Simon, L. S. (2009). Does it pay to be smart, attractive, or confident (or all three)? Relationships among general mental ability, physical attractiveness, core self-evaluations, and income. *Journal of Applied Psychology*, *94*, 742-755. http://dx.doi.org/10.1037/a0015497

*Judge, T. A., Ilies, R., & Zhang, Z. (2012). Genetic influences on core self-evaluations, job satisfaction, and work stress: A behavioral genetics mediated model. *Organizational Behavior and Human Decision Processes*, *117*, 208-220. http://dx.doi.org/10.1016/j.obhdp.2011.08.005

Judge, T. A., Locke, E. A., & Durham, C. C. (1997). The dispositional causes of job satisfaction: A core evaluations approach. *Research in Organizational Behavior*, *19*, 151-188.

Judge, T. A., Locke, E. A., Durham, C. C., & Kluger, A. N. (1998). Dispositional effects on job and life satisfaction: The role of core evaluations. *Journal of Applied Psychology*, *83*, 17-34. http://dx.doi.org/10.1037/0021-9010.83.1.17

*Judge, T. A., Van Vianen, A. E. M., & De Pater, I. E. (2004). Emotional stability, core self-evaluations, and job outcomes: A review of the evidence and an agenda for future research. *Human Performance*, *17*, 325-346. http://dx.doi.org/10.1207/s15327043hup1703_4

*Kacmar, K. M., Collins, B. J., Harris, K. J., & Judge, T. A. (2009). Core self-evaluations and job performance: The role of the perceived work environment. *Journal of Applied Psychology*, *94*, 1572-1580. http://dx.doi.org/10.1037/a0017498

*Kammeyer-Mueller, J. D., Judge, T. A., & Scott, B. A. (2009). The role of core self-evaluations in the coping process. *Journal of Applied Psychology*, *94*, 177-195. http://dx.doi.org/10.1037/a0013214

*Kim, T. -Y., Liden, R. C., Kim, S. P., & Lee, D. R. (2015). The interplay between follower core self-evaluation and transformational leadership: Effects on employee outcomes. *Journal of Business and Psychology, 30,* 345-355. http://dx.doi.org/10.1007/s10869-014-9364-7

*Kim, K., Oh, I. -S., Chiaburu, D. S., & Brown, K. S. (2012). Does positive perception of oneself boost learning motivation and performance? *International Journal of Selection and Assessment*, *20*, 257-271. http://dx.doi.org/10.1111/j.1468-2389.2012.00598.x

Klimoski, R. J. (1993). Predictor constructs and their measurement. In N. Schmitt and W. C. Borman (Eds.), *Personnel selection in organizations* (pp. 99-135). San Francisco, CA: Jossey-Bass.

*Kluemper, D. H. (2008). Trait emotional intelligence: The impact of core self-evaluations and social desirability. *Personality and Individual Differences*, *44*, 1402-1412. http://dx.doi.org/10.1016/j.paid.2007.12.008

*Kluemper, D. H., Mossholder, K. W., Ispas, D., Bing, M. N., Iliescu, D., & Ilie, A. (in press). When core self-evaluations influence employees' deviant reactions to abusive supervision: The moderating role of cognitive ability. *Journal of Business Ethics*. http://dx.doi.org/10.1007/s10551-018-3800-y

*Koumoundourou, G. A., Kounenou, K., & Siavara, E. (2012). Core self-evaluations, career decision self-efficacy, and vocational identity among Greek adolescents. *Journal of Career Development*, *39*, 269-286. http://dx.doi.org/10.1177/0894845310397361

*Kroeker, J. M. (2008). *The relationship between core self-evaluation traits and work family conflict* (Unpublished master's thesis). Emporia State University, Emporia, KS.

Lance, C. E., Butts, M. M., & Michels, L. C. (2006). The sources of four commonly reported cutoff criteria: What did they really say? *Organizational Research Methods*, *9*, 202-220. http://dx.doi.org/10.1177/1094428105284919

*Lavelle, M. G. (2006). *Supervisor effectiveness: A correlation study of supervisor self-assessment, and the management observation of effectiveness* (Unpublished doctoral dissertation). Capella University, Minneapolis, MN.

Law, K. S., Wong, C. S., & Mobley, W. H. (1998). Toward a taxonomy of multidimensional constructs. *Academy of Management Reviews*, *23*, 741-755. http://dx.doi.org/10.5465/amr.1998.1255636

Le, H., Schmidt, F. L., & Putka, D. (2009). The multifaceted nature of measurement artifacts and its implications for estimating construct-level relationships. *Organizational Research Methods, 12,* 165-200. http://dx.doi.org/10.1177/1094428107302900

Ledgerwood, A., & Shrout, P. E. (2011). The trade-off between accuracy and precision in latent variable models of mediation processes. *Journal of Personality and Social Psychology, 101,* 1174-1188. http://dx.doi.org/10.1037/a0024776

*Lee, K., & Duffy, M. K. (2019). A functional model of workplace envy and job performance: When do employees capitalize on envy by learning from envied targets? *Academy of Management Journal, 62,* 1085-1110. http://dx.doi.org/10.546/amj.2016.1202

*Liu, S., Wang, M., Piccolo, R., Zhan, Y., & Shi, J. (2008). *Core self-evaluation in China: Predicting job satisfaction and job performance*. Paper presented at the 23rd Annual Meeting of the Society for Industrial and Organizational Psychology, San Francisco, CA.

Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores.* Reading, MA: Addison-Mosley.

*Luria, G., & Torjman, A. (2009). Resources and coping with stressful events. *Journal of Organizational Behavior*, *30*, 685-707. http://dx.doi.org/10.1002/job.551

*Luthans, F., Avolio, B. J., Avey, J. B., & Norman, S. M. (2007). Positive psychological capital: Measurement and relationship with performance and satisfaction. *Personnel Psychology*, *60*, 541-572. http://dx.doi.org/10.1111/j.1744-6570.2007.00083.x

McAbee, S. T., Landis, R. S., & Burke, M. I. (2017). Inductive reasoning: The promise of big data. *Human Resource Management Review, 27,* 277-290. http://dx.doi.org/10.1016/j.hrmr.2016.08.005

McAbee, S. T., & Oswald, F. L. (2013). The criterion-related validity of personality measures for predicting GPA: A meta-analytic validity competition. *Psychological Assessment, 25,* 532-544. http://dx.doi.org/10.1037/a0031748

McDonald, R. P. (1999). *Test theory: A unified approach.* Mahwah, NJ: Erlbaum.

*McNall, L. A., & Mischel, J. S. (2011). A dispositional approach to work-school conflict and enrichment. *Journal of Business and Psychology*, *26*, 397-411. http://dx.doi.org/10.1007/s10869-010-9187-0

*McNall, L. A., Masuda, A. D., Shanock, L. R., & Nicklin, J. M. (2011). Interaction of core self-evaluations and perceived organizational support on work-to-family enrichment. *Journal of Psychology: Interdisciplinary and Applied*, *145*, 133-149. http://dx.doi.org/10.1080/00223980.2010.542506

*Meyers, R., & Houssemand, C. (2010). Socioprofessional and psychological variables that predict job finding. *Revue Europé enne de Psychologie Appliquée*, *60*, 201-219. http://dx.doi.org/j.erap.2009.11.004

*Mitchell, M. S., Greenbaum, R. L., Vogel, R. M., Mawritz, M. B., & Keating, D. J. (2019). Can you handle the pressure? The effect of performance pressure on stress appraisals, self-regulation, and behavior. *Academy of Management Journal, 62,* 531-552. http://dx.doi.org/10.5465/amj.2016.0646

*Mol, S. T., Born, M. Ph., Willemsen, M. E., van der Molen, H. T., & Derous, E. (2009). When selection ratios are high: Predicting the expatriation willingness of prospective domestic entry-level job applications. *Human Performance*, *22*, 1-22. http://dx.doi.org/10.1080/08959280802540437*

Moazami-Goodarze, A., Nurumi, J. E., Mauno, S., & Rantanen, J. (2015). Cross-lagged relations between work-family enrichment, vigor at work, and core self-evaluations: A three-wave study. *Journal of Business and Psychology, 30,* 473-482. http://dx.doi.org/10.1007/s10869-014-9376-3

*Murphy, W. M. (2011). From e-mentoring to blended mentoring: Increasing students' developmental initiation and mentors' satisfaction. *Academy of Management Learning & Education*, *10*, 606-622. http://dx.doi.org/10.5465/amle.2010.0090

*Narayan, A. (2008). *Role of assigned team goals in the relationship between individual difference factors and self-set goals in a pre-team context* (Unpublished doctoral dissertation). Wright State University, Dayton, OH.

*Neureiter, M., & Traut-Mattausch, E. (2017). Two sides of the career resources coin: Career adaptability resources and the imposter phenomenon. *Journal of Vocational Behavior, 98,* 56-69. http://dx.doi.org/10.1016/j.jvb.2016.10.002

*Newman, D. A. (2004). *Is job dissatisfaction contagious? Simultaneous effects of social networks, task characteristics, and dispositions* (Unpublished doctoral dissertation). Pennsylvania State University, University Park, PA.

*Ng, T. W. H., & Feldman, D. C. (2010). Idiosyncratic deals and organizational commitment. *Journal of Vocational Behavior*, *76*, 419-427. http://dx.doi.org/10.1016/j.jvb.2009.10.006

*Nikolaou, I., & Georgiou, K. (2018). Fairness reactions to the employment interview. *Journal of Work and Organizational Psychology, 34,* 103-111. http://dx.doi.org/10.5093/jwop2018a13

*Nübold, A., Muck, P. M., & Maier, G. W. (2013). A new substitute for leadership? Followers' state core self-evaluations. *Leadership Quarterly, 24,* 29-44. http://dx.doi.org/10.1016/j.leaqua.2012.07.002

Nunnally, J. (1978). *Psychometric theory.* New York, NY: McGraw-Hill.

*O'Neil, T. A., McLarnon, M. J. W., Xiu, L., & Law, S. J. (2016). Core self-evaluations, perception of group potency, and job performance: The moderating role of individualism and collectivism cultural profiles. *Journal of Occupational and Organizational Psychology, 89,* 447-473. http://dx.doi.org/10.1111/joop.12135

*Oostrom, J. K., Born, M. P., Serlie, A. W., & Van Der Molen, H. T. (2010). Effects of individual differences on the perceived job relatedness of a cognitive ability test and a multimedia situational judgment tests. *International Journal of Selection and Assessment*, *18*, 394-406. http://dx.doi.org/10.1111/j.1468-2389.2010.00521.x

Oswald, F. L., & Ercan, S. (2013). *Illustrating heterogeneity in meta-analysis using a new-yet-simple plot.* Paper presented at the Society for Research Synthesis Methods, Boston, MA.

Oswald, F. L., Ercan, S., McAbee, S. T., Ock, J., & Shaw, A. (2015). Imperfect corrections or correct imperfections? Psychometric corrections in meta-analysis. *Industrial and Organizational Psychology, 8,* e1-e4. http://dx.doi.org/10.1017/iop.2015.17

*Oyler, J. D. (2007). *Core self-evaluation and job satisfaction in the role of organizational and community embeddedness* (Unpublished doctoral dissertation). Virginia Polytechnic Institute and State University, Blacksburg, VA.

*Parrish, K. R. (2016). *Applying self-determination theory to examine the role of core self-evaluations on the relationship between job satisfaction and job performance: A moderated-mediation model* (Unpublished doctoral dissertation). Chicago School of Professional Psychology, Chicago, IL.

Pearlman, K., Schmidt, F. L., & Hunter, J. E. (1980). Validity generalization results for tests used to predict job proficiency and training success in clerical occupations. *Journal of Applied Psychology, 65,* 373-406. http://dx.doi.org/10.1037/0021-9010.65.4.373

Peterson, R. A., & Kim, Y. (2012). On the relationship between coefficient alpha and composite reliability. *Journal of Applied Psychology, 98,* 194-198. http://dx.doi.org/10.1037/a0030767

Podsakoff, P. M., MacKenzie, S. B., Lee, J. –Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, *88*, 879-903. http://dx.doi.org/10.1037/0021-9010.88.5.879

Polanin, J. R., Hennessy, E. A., & Tanner-Smith, E. E. (2017). A review of meta-analysis packages in R. *Journal of Educational and Behavioral Statistics, 42,* 206-242. http://dx.doi.org/10.3102/1076998616674315

*Quigley, N. R. (2003). *The relationship between leader core self-evaluations, team feedback, leader efficacy, transformational leadership, team efficacy, team goals, team action and transition process, and team performance* (Unpublished doctoral dissertation). University of Maryland, College Park, MD.

*Ramirez, S. A. (2015). *Impulsive and premeditated counterproductive work behaviors and the moderating effects of self-monitoring and core self-evaluation* (Unpublished doctoral dissertation). North Carolina State University, Raleigh, NC.

Raykov, T. (1997). Scale reliability, Cronbach's coefficient alpha, and violations of essential tau-equivalence with fixed congeneric components. *Multivariate Behavioral Research, 32,* 329-353. http://dx.doi.org/10.1207/s15327906mbr3204_2

*Rey, L., & Extremera, N. (2015). Core self-evaluations, perceived stress and life satisfaction in Spanish young and middle-aged adults: An examination of mediation and moderation effects. *Social Indicators Research, 120,* 515-524. http://dx.doi.org/10.1007/s11205-014-0601-2

*Rey, L., Extremera, N., & Peláez-Fernández, M. A. (2016). Linking social support to psychological distress in the unemployed: The moderating role of core self-evaluations. *Social Indicators Research, 127,* 435-445. http://dx.doi.org/10.1007/s11205-015-0958-x

*Ronen, E. (2008). *Transfer of e-learning in the workplace: The effects of trainee characteristics and contextual factors* (Unpublished doctoral dissertation). Illinois Institute of Technology, Chicago, IL.

*Rosopa, P. J., & Schroeder, A. N. (2009). Core self-evaluations interact with cognitive ability to predict academic achievement. *Personality and Individual Differences*, *47*, 1003-1006. http://dx.doi.org/j.paid.2009.07.028

Scherer, R., & Teo, T. (2020). A tutorial on the meta-analytic structural equation modeling of reliability coefficients. *Psychological Methods, 25,* 747-775. http://dx.doi.org/10.1037/met0000261

Schmidt, F. L., & Hunter, J. E. (1999). Comparison of three meta-analysis methods revisited: An analysis of Johnson, Mullen, and Salas (1995). *Journal of Applied Psychology, 84,* 144-148. http://dx.doi.org/10.1037/0021-9010.84.1.144

Schmidt, F. L., & Hunter, J. E. (2015). *Methods of meta-analysis: Correcting error and bias in research findings* (3rd ed.). Thousand Oaks, CA: Sage.

Schmidt, F. L., Oh, I. -S., & Hayes, T. L. (2009). Fixed- versus random-effects models in meta-analysis: Model properties and an empirical comparison of differences in results. *British Journal of Mathematical and Statistical Psychology*, *62*, 97-128. http://dx.doi.org/10.1348/000711007X255327

Schmitt, N. (2004). Beyond the Big Five: Increases in understanding and practical utility. *Human Performance, 17,* 347-357. http://dx.doi.org/10.1207/s1532704hup1703_5

Schmitt, N., & Landy, F. J. (1993). The concept of validity. In N. Schmitt & W. C. Borman (Eds.), *Personnel selection in organizations* (pp. 275-309). San Francisco, CA: Jossey-Bass.

*Sears, G. J., & Hackett, R. D. (2011). The influence of role definition and affect in LMX: A process perspective on the personality – LMX relationship. *Journal of Occupational and Organizational Psychology*, *84*, 544-564. http://dx.doi.org/10.1348/096317910X492081

*Sears, G. J., & Holmvall, C. M. (2010). The joint influence of supervisor and subordinate emotional intelligence on leader-member exchange. *Journal of Business and Psychology*, *25*, 593-605. http://dx.doi.org/10.1007/s10869-009-9152-y

*Shamas, T. M. (2009). *Influence of pre-test information on applicant perceptions: An attributional approach* (Unpublished doctoral dissertation) Alliant International University, San Diego, CA.

*Shantz, A., & Booth, J. E. (2014). Service employees and self-verification: The roles of occupational stigma consciousness and core self-evaluations. *Human Relations, 67,* 1439-1465. http://dx.doi.org/10.1177/0018726713519280

*Smith, J. N. (2017). *The impact of core self-evaluations and team conflict on individual burnout*. Unpublished doctoral dissertation. University of Calgary, Calgary, AB.

*Soane, E., Booth, J. E., Kerstin, A., Shantz, A., & Bailey, C. (2018). Deadly combinations: how leadership contexts undermine the activation and enactment of followers' high core self-evaluations in performance. *European Journal of Work and Organizational Psychology, 27,* 297-309. http://dx.doi.org/10.1080/1359432X.2018.1444602

*Song, Z., Chon, K., Ding, G., & Gu, C. (2015). Impact of organizational socialization tactics on newcomer job satisfaction and engagement: Core self-evaluations as moderators. *International Journal of Hospitality Management, 46,* 180-189. http://dx.doi.org/10.1016/j.ijhm.2015.02.006

Spector, P. E., Rogelberg, S. G., Ryan, A. M., Schmitt, N., & Zedeck, S. (2014). Moving the pendulum back to the middle: Reflections on the introduction to the inductive research special issue of Journal of Business and Psychology. *Journal of Business and Psychology, 29,* 499-502. http://dx.doi.org/10.1007/s10869-014-9372-7

*Spitzmuller, C. (2003). *Investigating mediators of the core self-evaluations – job satisfaction relationship* (Unpublished doctoral dissertation). Bowling Green State University, Bowling Green, OH.

*Steele-Johnson, D., & Narayan, A. (2007). *Goal orientations, core self-evaluations, and self-efficacy in a* *team context*. Paper presented at the 22nd Annual Meeting of the Society for Industrial and Organizational Psychology, New York, NY.

*Stein, D., & Grant, A. M. (2014). Disentangling the relationships among self-reflection, insight, and subjective well-being: The role of dysfunctional attitudes and core self-evaluations. *Journal of Psychology, 148,* 505-522. http://dx.doi.org/10.1080/00223980.2013.810128

*Stumpp, T., Hülsheger, U. R., Muck, P. M., & Maier, G. W. (2009). Expanding the link between core self-evaluations and affective job attitudes. *European Journal of Work and Organizational Psychology*, *18*, 148-166. http://dx.doi.org/10.1080/13594320802281110

*Stumpp, T., Muck, P. M., Hülsheger, U. R., Judge, T. A., & Maier, G. W. (2010). Core self-evaluations in Germany: Validation of a German measure and its relationships with career success. *Applied Psychology: An International Review*, *59*, 674-700. http://dx.doi.org/10.1111/j.1464-0597.2010.00422.x

*Sumanth, J. J., & Cable, D. M. (2011). Status and organizational entry: How organizational and individual status affect justice perceptions of hiring systems. *Personnel Psychology*, *64*, 963-1000. http://dx.doi.org/10.1111/j.1744-6570.2011.01233.x

*Sun, P., & Jiang, H. (2017). Psychometric properties of the Chinese version of Core Self-Evaluations Scale. *Current Psychology, 36,* 297-303. http://dx.doi.org/10.1007/s12144-016-9418-2

*Sweet, K. M. (2012). *Dysfunctional leader-member exchanges: Reexamining leader and follower personality and relationship formation* (Unpublished doctoral dissertation). University of Houston, Houston, TX.

*Swody, C. A. (2006). *Work-family enrichment: The role of work eustress in spillover of positive psychological states* (Unpublished doctoral dissertation). University of Connecticut, Storrs-Mansfield, CT.

*Tasa, K., Sears, G. J., Schat, A. C. H. (2011). Personality and teamwork behavior in context: The cross-level moderating role of collective efficacy. *Journal of Organizational Behavior*, *32*, 65-85. http://dx.doi.org/10.1002/job.680

*Taylor, C. W. (2004). *The relationship between core self-evaluations and job satisfaction: Finding the mediating variables* (Unpublished doctoral dissertation). University of Alabama, Tuscaloosa, AL.

*Tepper, B. J., Carr, J. C., Breaux, D. M., Geider, S., Hu, C., & Hua, W. (2009). Abusive supervision, intentions to quit, and employees' workplace deviance: A power/dependence analysis. *Organizational Behavior and Human Decision Processes*, *109*, 156-167. http://dx.doi.org/10.1016/j.obhdp.2009.03.004

*Tews, M. J., Michel, J. W., & Noe, R. A. (2017). Does fun promote learning? The relationship between fun in the workplace and informal learning. *Journal of Vocational Behavior, 98,* 46-55. http://dx.doi.org/10.1016/j.jvb.2016.09.006

Thompson, B. (2003). *Score reliability: Contemporary thinking on reliability issues.* Thousand Oaks, CA: Sage.

*Thompson, C. N. (2008). *Personal characteristics and the impact of transformational leadership behaviors on follower outcomes* (Unpublished doctoral dissertation). Wright State University, Dayton, OH.

*Tillman, C. J. (2011). *Character, conditions, and cognitions: The role of personality, climate, intensity, and moral disengagement in the unethical decision-making process* (Unpublished doctoral dissertation). University of Alabama, Tuscaloosa, AL.

*Triana, M. C. (2008). *Are virtual teams more just? An investigation of how reducing social categorization can increase female participation in male-dominated teams* (Unpublished doctoral dissertation). Texas A & M University, College Station, TX.

*Tsaousis, I., Nikolaou, I., Serdaris, N., & Judge, T. A. (2007). Do the core self-evaluations moderate the relationship between subjective well-being and physical and psychological health? *Personality and Individual Differences*, *42*, 1441-1452. http://dx.doi.org/10.1016/j.paid.2006.10.025

Vacha-Haase, T. (1998). Reliability generalization: Exploring variance in measurement error affecting score reliability across studies. *Educational and Psychological Measurement*, *58*, 6-20. http://dx.doi.org/10.1177/0013164498058001002

*van Doorn, R. R. A., Lang, J. W. B., & Weijters, T. (2010). Self-reported cognitive failures: A core self-evaluation? *Personality and Individual Differences*, *49*, 717-722. http://dx.doi.org/10.1016/j.paid.2010.06.013

*Van Hoye, G., Saks, A. M., Lievens, F., & Weijters, B. (2015). Development and test of an integrative model of job search behavior. *European Journal of Work and Organizational Behavior, 24,* 544-559. http://dx.doi.org/10.1080/1359432X.2014.964214

*Venz, R. A., & Gardiner, E. (2017). It pays to be well-connected. The moderating role of networking ability on the relationship between core self-evaluations and income. *Personality and Individual Differences, 110,* 85-89. http://dx.doi.org/10.1016/j.paid.2017.01.032

*Volmer, J. (2015). Followers' daily reactions to social conflicts with supervisors: The moderating role of core self-evaluations and procedural justice perceptions. *Leadership Quarterly, 26,* 719-731. http://dx.doi.org/10.1016/j.leaqua.2015.01.005

*Volmer, J., Koch, I. K., & Wolff, C. (2019). Illuminating the 'dark core': Mapping global versus specific sources of variance across multiple measures of the dark triad. *Personality and Individual Differences, 145,* 97-102. http://dx.doi.org/10.1016/j.paid.2019.03.024

*Wagstaff, M. F., del Carmen Triana, M., Kim, S., & Al-Riyami, S. (2015). Responses to discrimination: Relationships between social support seeking, core self-evaluations, and withdrawal behaviors. *Human Resource Management, 54,* 673-687. http://dx.doi.org/10.1002/hrm.21634

*Wanberg, C. R., Glomb, T. M., Song, Z., & Sorenson, S. (2005). Job-search persistence during unemployment: A 10-wave longitudinal study. *Journal of Applied Psychology*, *90*, 411-430. http://dx.doi.org/10.1037/0021-9010.90.3.411

*Wang, G., Van Iddekinge, C. H., Zhang, L., & Bishoff, J. (2019). Meta-analytic and primary investigations of the role of followers in ratings of leadership behavior in organizations. *Journal of Applied Psychology, 104,* 70-106. http://dx.doi.org/10.1037/apl0000345

*Wang, Z., & Xu, H. (2019). When and for whom ethical leadership is more effective in eliciting work meaningfulness and positive attitudes: The moderating roles of core self-evaluation and perceived organizational support. *Journal of Business Ethics, 156,* 919-940. http://dx.doi.org/10.1007/s10551-017-3563-x

*Weider, N. W. (2012). *Dissecting implicit leadership theories: A generalizability analysis* (Unpublished doctoral dissertation). Wright State University, Dayton, OH.

*Westring, A. F., & Ryan, A. M. (2010). Personality and inter-role conflict and enrichment: Investigating the mediating role of support. *Human Relations*, *63*, 1815-1834. http://dx.doi.org/10.1177/0018726710371236

*Westring, A. F., & Ryan, A. M. (2011). Anticipated work-family conflict: A construct investigation. *Journal of Vocational Behavior*, *79*, 596-610. http://dx.doi.org/10.1016/j.jvb.2011.02.004

Whitener, E. M. (1990). Confusion of confidence intervals and credibility intervals in meta-analysis. *Journal of Applied Psychology, 75,* 315-321. http://dx.doi.org/10.1037/0021-9010.75.3.315

*Williams, K. Z. (2012). *Does practice make perfect? Effects of practice and coaching on interview performance* (Unpublished doctoral dissertation). Clemson University, Clemson, SC.

*Woolum, A., Foulk, T., Lanaj, K., & Erez, A. (2017). Rude color glasses: The contaminating effects of witnessed morning rudeness on perceptions and behaviors throughout the workday. *Journal of Applied Psychology, 102,* 1658-1672. http://dx.doi.org/10.1037/apl0000247

*Wu, T., & Hu, C. (2009). Abusive supervision and employee emotional exhaustion. *Group & Organizational Management*, *34*, 143-169. http://dx.doi.org/10.1177/1059601108331217

*Xu, E., Huang, X., & Robinson, S. L. (2017). When self-view is at stake: Responses to ostracism through the lens of self-verification theory. *Journal of Management*, *43,* 2281-2302. http://dx.doi.org/10.1177/0149206314567779

*Xu, X-. M., & Yu, K. (2019). When core self-evaluation leads to career adaptability: Effects of ethical leadership and implications for citizenship behavior. *Journal of Psychology: Interdisciplinary and Applied, 153,* 463-477. http://dx.doi.org/10.1080/00223980.2018.1564724

*Yuan, Z., Li, Y., & Lin, J. (2014). Linking challenge and hindrance stress to safety performance: The moderating effect of core self-evaluation. *Personality and Individual Differences, 68,* 154-159. http://dx.doi.org/10.1016/j.paid.2014.04.025

*Zacher, H. (2016). Within-person relationships between daily individual and job characteristics and daily manifestations of career adaptability.

Journal of Vocational Behavior, 92, 105-115. http://dx.doi.org/10.1016/j.jvb.2015.11.013

*Zacher, H. (2019). Why do we act as old as we feel? The role of occupational future time perspective and core self-evaluations in the relationship between subjective age and job crafting behavior. *European Journal of Work and Organizational Psychology, 28,* 831-844. http://dx.doi.org/10.1080/1359432X.2019.1677609

*Zare, M. (2017). *A multilevel approach in examining abusive supervision: The role of core self-evaluation and identity threat* (Unpublished doctoral dissertation). New Mexico State University, Las Cruces, NM.

*Zenger, M., Körner, A., Maier, G. W., Hinz, A., & Stöbel-Richter, Y., Brähler, E., & Hilbert, A. (2015). The core self-evaluations scale: Psychometric properties of the German version in a representative sample. *Journal of Personality Assessment, 97,* 310-318. http://dx.doi.org/10.1080/00223891.2014.989367

*Zhang, H., Kwan, H. K., Zhang, X., & Wu, L. (2014). High core self-evaluators maintain creativity: A motivational model of abusive supervision. *Journal of Management, 40,* 1151-1174. http://dx.doi.org/10.1177/0149206312460681

*Zhang, Z., & Peterson, S. J. (2011). Advice networks in teams: The role of transformational leadership and members' core self-evaluations. *Journal of Applied Psychology*, *96*, 1004-1017. http://dx.doi.org/10.1037/a0023254

*Zhu, J., Wanberg, C. R., Harrison, D. A., & Diehn, E. W. (2016). Ups and downs of the expatriate experience? Understanding work adjustment trajectories and career outcomes. *Journal of Applied Psychology, 101,* 549-568. http://dx.doi.org/10.1037/apl0000073

Citation:

Ock, J., McAbee, S. T., Ercan, S., Shaw, A., & Oswald, F. L. (2021). Reliability generalization analysis of the Core Self-Evaluations Scale *Practical Assessment, Research & Evaluation*, 26(6). Available online: https://scholarworks.umass.edu/pare/vol26/iss1/6/

Corresponding Author:

Jisoo Ock
Department of Business Administration
Pusan National University
Busan, South Korea

Email: jisoo.ock [at] pusan.ac.kr