



A comprehensive analysis of the psychometric properties of the contingencies of self-worth scale (CSWS)

Enrico Perinelli¹ · Guido Alessandri² · Michele Vecchione³ · David Mancini²

© The Author(s) 2020

Abstract

The Contingencies of Self-Worth Scale (CSWS) is a widely used personality self-report questionnaire developed for measuring the domains in which self-esteem is sustained by successes and achievements as well as threatened by obstacles and failures. Two studies ($N_{\text{study1}} = 453$, $N_{\text{study2}} = 293$) aimed to further refine our knowledge of its psychometric properties. Results attested that, at the first-order level, the originally hypothesized seven-factor model proved to be the best-fitting one, but the inclusion of a method factor significantly improved the fit to the data. At the second-order level, the model with two higher-order variables representing *private sphere* and *public sphere* of CSW fit better than alternative models. Finally, there was evidence that first- and second-order domains had a good degree of construct and discriminant validity. Overall, these studies provided a step forward in refining the psychometric structure of the CSWS.

Keywords Contingencies of self-worth scale · Self-esteem · Method factor · Higher-order model · Second-order model

People differ about the domains they regard as relevant for their self-worth. Some students, for example, might base their self-esteem on scholastic activities, leading to a strong relationship of school achievement with self-esteem. Other students, instead, might invest more in social relationships or physical appearance. For them, the relationship between school achievement and self-esteem is expected to be very weak or nonsignificant. Indeed, according to James (1890), feelings of personal worth are entirely dependent on one's objectives in the world and thus on one's achievements in

valued domains. Drawing upon James's intuition, Crocker and Wolfe (2001) introduced the construct of contingencies of self-worth and defined them as "the domains in which self-esteem is bolstered by successes and achievements, and threatened by setbacks and failures" (Crocker et al. 2003, p. 894). Accordingly, when people are faced with negative events in a specific valued domain, defensive responses or clear reductions to self-esteem levels are expected (see Crocker et al. 2003).

To assess the hypothesized seven sources of self-esteem and to test their model, Crocker et al. (2003) developed the 35-item Contingencies of Self-Worth Scale (CSWS). This scale assesses people's perceived sense that their own judgments of self-worth are influenced by their behavior or by the outcomes reached in key domains of experience. Since its introduction, this scale has been widely adopted (with 1158 references in Google Scholar as of March 6, 2020), has been adapted in seven cultures (Self and Social Motivation Lab 2015), and has consistently shown good psychometric properties in several earlier empirical studies (see Crocker and Luhtanen 2003; Crocker et al. 2003; Luhtanen and Crocker, 2005; Park et al. 2007).

However, a review of the literature revealed that the hypothesized 7-factor model supported by Crocker et al. (2003) failed to show a consistently good fit to the data in several successive studies (i.e., Bentea 2016; Kempnaers

Enrico Perinelli and Guido Alessandri should be considered co-first authors. They contributed equally to this work and the order of their names was arbitrary.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s12144-020-01007-5>) contains supplementary material, which is available to authorized users.

✉ Guido Alessandri
guido.alessandri@uniroma1.it

¹ Department of Psychology and Cognitive Science, University of Trento, Trento, Italy

² Department of Psychology, Sapienza University of Rome, Via dei Marsi 78, 00185 Rome, Italy

³ Department of Social and Developmental Psychology, Sapienza University of Rome, Rome, Italy

et al. 2014; Maricuțoiu et al. 2012). Thus, we hypothesized that one likely source of misfit is represented by the inclusion of positively and negatively worded items. While the presence of positively and negatively worded items has been often recommended to prevent response styles (Baumgartner and Steenkamp 2001; Paulhus 1991; Podsakoff et al. 2012), their presence has been documented to lead to several issues such as ingenerating artificial deviation from the hypothesized scale structure (Alessandri et al. 2011). We deemed this hypothesis compelling given that similar issues have been attested for other instruments used to assess individuals' self-esteem (see Alessandri et al. 2015).

We conducted two studies with the aim of providing a full evaluation of the psychometric properties of the CSWS in terms of structural validity and external correlates. Furthermore, we introduced and investigated the validity of a new bifactorial structure of the scale that controls for method effects due to item wording, thus allowing for a purer measurement of each of the seven domains. Moreover, we conducted an in-depth investigation of the second-order structure of the CSWS, examining its congruency with the theory suggesting that the seven first-order domains reflect the two basic typologies of intra-individual vs. inter-individual evaluations (see Crocker et al. 2003). We conducted our analyses on two samples, on a combined number of about 500 individuals, using cross-sectional, lagged, and intensive data. In this way, we hope to contribute to the growing literature regarding the CSWS by offering a detailed overview of its psychometric properties.

Contingencies of Self-Worth

According to Crocker and Wolfe (2001), the contingencies of self-worth (CSW) are the domains on which people's self-esteem depend, so that progresses or successes in these domains boost self-esteem, and failures or setbacks lead to reductions in self-esteem. The CSW were theoretically derived on the assumption that one's self-regard is based, at least in part, on a set of valued *intrapersonal* and *interpersonal* domains (see Crocker et al. 2003; Crocker and Wolfe 2001). The CSW model of self-esteem focuses on the domains on which people stake their self-worth rather than on differences between people in whether they have contingent self-esteem or not. The authors of the model (Crocker and Wolfe 2001) selected seven domains, starting from the most known and widely adopted domains of *competencies* and *social approval* and further expanding them to include other common sources of self-esteem. Overall, these seven theoretically-derived CSW domains are clustered into two broad domains.

The first is the "intrapersonal domain" which includes the subdomains of *virtue* and *God's love*. Most individuals base their sense of worthiness on being coherent to a moral code, which leads to the feeling of being a worthwhile person (Solomon et al. 1991). Therefore, *virtue* has been posited as one of the seven domains of the CSW. The domain of *God's love* reflects the fact that one's belief of being loved and unique in God's eyes might have a positive effect on self-esteem, especially for religious individuals (Benson and Spilka 1973; Blaine and Crocker 1995; Spilka et al. 1985).

The second is the "interpersonal domain" which includes the five subdomains of *academic competence*, *competition*, *approval from others*, *appearance*, and *family support*. Peoples' self-evaluation of the degree of CSW in these areas are conceived to be correlated but distinct from general self-esteem (i.e., judgments of overall self-worth; see Crocker et al. 2003; Crocker and Wolfe 2001). The domain of competence, precisely of *academic competence*, has been selected because many previous studies (e.g., Coopersmith 1967; Demo and Parker 1987; Harter 1986; Hoge et al. 1990; Rosenberg et al. 1995; Richman et al. 1987) found that academic outcomes are related to global self-esteem. The domain of *competition* is linked with self-esteem because people might base their self-esteem not only on being competent but rather on being better than others (Cross & Madson, 1997; Josephs, Markus, & Tafarodi, 1992). The *approval from others* domain regards what people believe others think of them, and it is taken into account because this belief plays an important role for global self-esteem (e.g., Cooley 1902; Coopersmith 1967; Harter 1986; Mead 1934; Shrauger and Schoeneman 1979; Wylie 1979). Since the affection of close others might be particularly important to self-esteem, and perceived approval or love from family members is related to self-worth, *family support* is one of the seven domains of the contingencies of self-worth (Harter 1986). Finally, the domain of *appearance* has been included because it is one of the strongest predictors of global self-esteem among adolescents (Harter 1986).

Present Contribution

Given the above considerations, as following we introduce our studies, which aim at investigating some psychometric characteristics of the CSW scale. In Study 1, we conduct a thorough evaluation of the internal structure of the CSW scale. In particular, we would investigate the fit and the comparison of several first- and second-order competing models, in order to find the best solutions. Then, in Study 2, we try to mimic Study 1 findings in terms of internal structure, and we also investigate the relationship of CSW dimensions (first- and

second-order) in regards of external variables of interest, in order to provide a better understanding of the nomological network of CSW. In Supplementary Material Section, we reported *Mplus* scripts for running all analyses.

Study 1

We designed this first study to conduct an in-depth evaluation of the structure of the CSWS at both the first- and second-order level using a large sample of university students. To this aim, we will investigate the competitive fit of alternative first-order structures that have been proposed in the literature. Then we will move to the second-order level, where the dimensionality of the scale has been seldom analyzed. This is surprising given the growing interest of researchers in investigating the higher-order structure of different constructs and scales (e.g., Isiordia et al. 2017). Moreover, it is noteworthy to say that a higher-order structure was initially hypothesized by Crocker et al. (2003) several years ago, but its robustness has never been tested empirically. Thus, we aimed at comparing different second-order models in order to find the best fitting one. In the following section, we describe the previous literature and the rationale of our goals in more detail.

Aim 1: Analysis of the First-Order Structure

For evaluating the structural validity of the CSW scale at the first-order level, we compared its theoretical model (positing seven correlated factors representing the seven theoretical CSW domains) with a set of alternative models proposed in the literature or representing extensions of existing models. Each of these models offer a very different interpretation of CSWS structure, and are all presented in Fig. 1. Model 1 is the theoretical model described above in which the seven domains of CSW were correlated. The following three models were introduced by Crocker et al. (2003), and they have since been investigated in previous studies on the dimensionality of the CSWS. Model 2 is a one-factor model with all items loading on a single general dimension of contingent self-worth. Model 3 posits two latent dimensions representing internal and external contingencies (Crocker et al. 2003). Accordingly, items assessing *God's love* and *virtues* were loaded on the first factor (internal contingencies), whereas items assessing the remaining five dimensions were loaded on the second dimension (external contingencies). Model 4 is a three-factor model including the dimensions of “Self-esteem goes up”, “Self-esteem goes down”, and “Self-esteem depends” (see Crocker et al. 2003). We expected Model 1 to show a better fit than Models 2–4. However, given the presence of negatively worded items, we further investigated the presence of method

variance by explicitly introducing a negative method factor in the model. Accordingly, Model 5 was a bifactor model suited to investigate the impact of method effect on scale dimensionality. This model used a bifactor approach (Reise 2012) for modeling method effect along with substantive factors (see Alessandri et al. 2010, 2011, 2015) and thus allowed us to disentangle true CSW variance from method variance ingenerated by the item formulation and keying. This model refines our understanding of item behavior and represents an improved version of Model 1 (rather than an alternative model).

More in detail, Model 5 used a version of the *correlated trait–correlated method minus one* framework (Eid et al. 2003), where “method” refers to the direction of item wording (keying). The *M - 1* approach was widely suggested in presence of method factors, given that it has been showed to reduce model complexity, to improve easiness of parameter interpretation, and to enhance model estimation (see Geiser et al. 2012; Geiser and Lockhart 2012; Hintz et al. 2019). The *positive wording method* was chosen as the comparison standard, given that the CSWS has a lower number of negatively worded items (i.e., 7) than positively worded items. Therefore, in this way, a fewer number of parameters would be estimated. As a consequence, we modeled an orthogonal (i.e., not correlated with specific or substantive factors) latent factor that accounted for the residual variance shared by negatively worded items. This model is essentially a bifactor model with correlated factors (given the supposed second-order structure of the scale and the presence of a method factor; see Rindskopf and Rose 1988), which accounted for the covariation among CSWS items in terms of (a) the seven broad correlated general factors attested by the literature and (b) one method factor associated with negatively worded items (see Fig. 1, Model 5).

Aim 2: Analysis of the Higher-Order Factor Structure

The second aim, which is the more novel aim of this study, was to refine our understanding of the higher-order structure of the CSWS. As stated in the premise, the seven CSW dimensions were conceived as reflecting two higher-order domains of CSW, namely the *intrapersonal* and the *interpersonal*. Crocker et al. (2003) explicitly noted that “another possibility for describing the relations among contingencies is that the seven distinct factors are organized within two higher order factors, one for external contingencies and one for internal contingencies” (Crocker et al. 2003, p. 896). However, no previous studies have explicitly explored this possibility. Assuming the existence of the two higher-order factors of intraindividual and interindividual contingencies, one would expect to find a pattern of intercorrelations among

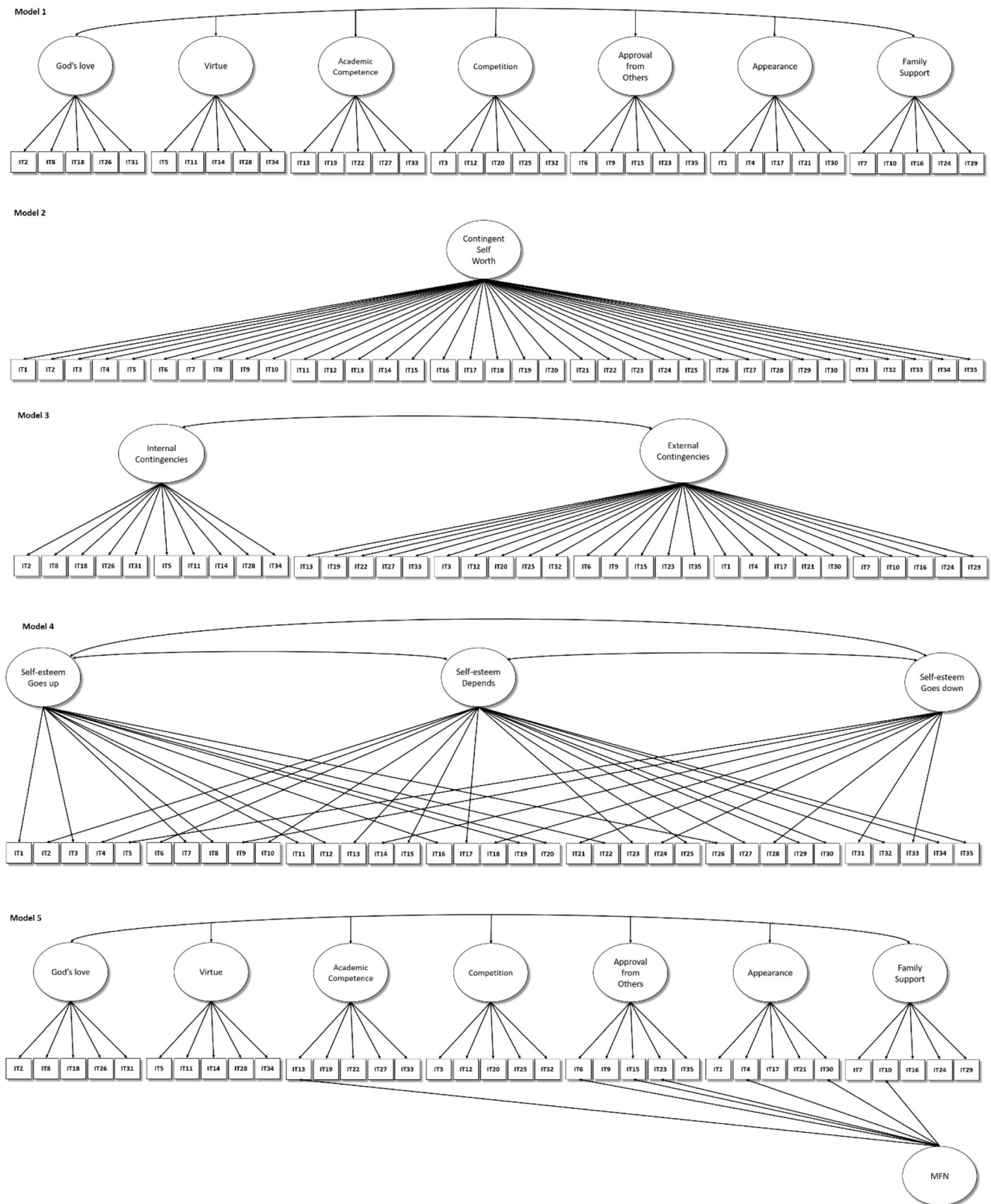


Fig. 1 The five alternative first-order models for the Contingencies of Self-Worth Scale (CSWS). In each model, residual variances for the observed variables were omitted for sake of clarity. IT = Item; MFN = Method Factor associated with Negatively worded items

the seven CSW first-order dimensions compatible with that hypothesis (Brown 2015). If this pattern holds, we would expect that a model positing two higher-order factors representing the intrapersonal and interpersonal dimensions (loaded by *God's love* and *virtue* and by the remaining five dimensions, respectively; see Fig. 1A in Supplementary Material, Model 6) would fit the data equally well (but it is more parsimonious in terms of free parameters) than a model with seven correlated factors.

An alternative interpretation of the second-order structure of the CSWS has been empirically derived by Stefanone et al. (2011). These authors conducted an exploratory factor analysis on the CSWS and derived two factors. The first factor was loaded by the dimensions of *family support*, *virtue*, and *God's love*. Given that the content of this factor appeared to reflect elements related to more traditional, personal domains, it was labeled *private sphere* of CSW. The second factor was loaded on the dimension of *approval from others*, *appearance*, and *competition* and was labeled *public sphere* of CSW. *Academic competence* cross-loaded on both factors (see Fig. 1A in Supplementary Material, Model 7).

We compared these two alternative second-order models against each other and against another alternative model positing a single higher-order factor representing a general CSW latent factor (see Fig. 1A in Supplementary Material, Model 8). This factor was specified as loaded by all the first-order dimensions, representing the alternative hypothesis that CSW blends into a single second-order dimension.

Method

Participants and Procedures

The present study was based on a convenience sample including 453 Italian sophomore students (72.6% females), all Caucasian. The participation was voluntary and no exclusion criteria were applied (nor any outlying participant was detected or excluded). Participants were recruited by a team of researchers and agreed to complete a set of questionnaires administered by using an online platform directly at their homes. For their participation in the study, participants were offered feedback about their psychological profile. The age ranged from 17 to 58 years with a mean of 21.52 years ($SD = 4.25$).

Measures

The Contingencies of Self-Worth Scale (CSWS; Crocker et al. 2003) comprises 35 items, five for each of the seven types of contingencies described by the CSW model. Seven items are negatively worded and have been reversed before computing the total score. Participants are requested to evaluate each item using a 7-point Likert scale (from 1 = *Strongly Disagree* to

7 = *Strongly Agree*). Reliability for each of the seven subscales was widely acceptable (all Cronbach's alphas and omegas were above .75; see Table 1A in Supplementary Material).

Statistical Analyses

Given the ordinal nature of the data, models were tested with the statistical software *Mplus* 8 (Muthén and Muthén 1998–2017) using the weighted least squares mean- and variance-adjusted (WLSMV) estimation method. This estimator provides weighted least squares parameter estimates using a diagonal weight matrix and robust standard errors and a mean- and covariance-adjusted χ^2 test statistic (see Finney and DiStefano 2013), and it is especially suited for models with categorical variables (DiStefano et al. 2018). Model fit was evaluated, in addition to the previously presented χ^2 , by looking at values of the comparative fit index (CFI), Tucker-Lewis index (TLI), and the root-mean-square error of approximation (RMSEA). We accepted values of CFI and TLI higher than .95, and RMSEA values lower than 0.06 (Hu and Bentler 1999) or with a 90% upper limit lower than 0.10 (see Kline 2016). Hence, as a general rule, we deemed a model as fitting the data when showing values of approximate fit indices (CFI, TLI, RMSEA) within the aforementioned thresholds.

For comparing all models, given that AIC is not computed by *Mplus* when using a categorical estimator, we adapted a version of the AIC proposed by Yamaoka et al. (1978) for weighted least squares models (see also Banks and Joyner 2017; Saleh 2014), computed as follows: $AIC = n \cdot \ln(\hat{f}_{min}) + 2p + 1$, where n is the sample size, \ln is the natural log, \hat{f}_{min} is the minimum value of the WLSMV fitting function (final iteration), and p is the number of free parameters (see Banks and Joyner 2017, p. 38, eq. 10). AIC rewards goodness of fit and includes a penalty that is an increasing function of the number of parameters estimated. We rescaled AIC values according to the following formula (Burnham and Anderson 2004): $\Delta AIC = AIC_i - AIC_{min}$, where AIC_{min} is the minimum of the observed AIC values (among the i competitive models). This transformation forces the best model to have $\Delta AIC = 0$ while the rest of the models have positive values. Accordingly, a model that differs less than $\Delta AIC = 2$ from the best fitting model in a specific dataset is said to be "strongly supported by the evidence." If the difference lies between $4 \leq \Delta AIC \leq 7$ there is considerably less support, whereas models with $\Delta AIC > 10$ have essentially no support (Burnham and Anderson 2004, p. 271).

In sum, the structural (internal) validity of a model (either first-order or higher-order) is attested whether a model shows (a) good fit indices and (b) a better fit than competing models.

Table 1 Results from First- and Second-Order Confirmatory Factor Analyses

Study	Model	WLSMV-based χ^2	df	CFI	TLI	RMSEA (CI 90%)	AIC	Δ AIC
Study 1	First-order models							
	Model 1: Theoretical 7-domain model	1585.30	539	.963	0.960	0.065 (0.062, 0.069)	811.07	43.92
	Model 2: One-factor model	10,338.92	560	.659	0.637	0.196 (0.193, 0.200)	960.61	193.46
	Model 3: Internal vs external	4593.10	559	.859	0.850	0.126 (0.123, 0.130)	1511.63	744.48
	Model 4: Three-factor model (up, down, depend)	No convergence	–	–	–	–	–	–
	Model 5: 7-domain model + MFN	1437.29	532	.968	0.965	0.061 (0.057, 0.065)	767.15	0
	Second-order models							
	Model 6: Crocker et al.'s (2003) hypothesized higher-order model	1867.67	545	.954	0.950	0.073 (0.070, 0.077)	1004.23	123.47
Study 2	Model 7: Stefanone et al.'s (2011) hypothesized higher-order model	1486.78	544	.967	0.964	0.062 (0.058, 0.066)	880.76	0
	Model 8: Higher-order one-factor model	2005.66	546	.949	0.944	0.077 (0.073, 0.080)	1061.85	181.09
	First-order models							
	Model 1: Theoretical 7-domain model	1276.488	539	.962	0.958	0.068 (0.064, 0.073)	776.62	9.37
	Model 2: One-factor model	6907.448	560	.671	0.651	0.197 (0.193, 0.201)	1469.87	702.61
	Model 3: Internal vs external	3256.784	559	.860	0.851	0.128 (0.124, 0.133)	1181.84	414.58
	Model 4: Three-factor model (up, down, depend)	No convergence	–	–	–	–	–	–
	Model 5: 7-domain model + MFN	1209.323	532	.965	0.961	0.066 (0.061, 0.071)	767.26	0
Second-order models								
Model 6: Crocker et al.'s (2003) hypothesized higher-order model	1476.985	545	.952	0.947	0.076 (0.072, 0.081)	893.18	61.69	
Model 7: Stefanone et al.'s (2011) hypothesized higher-order model	1255.96	544	.963	0.960	0.067 (0.062, 0.072)	831.49	0	
Model 8: Higher-order one-factor model	1549.223	546	.948	0.943	0.079 (0.075, 0.084)	919.35	87.86	

Note. Best fitting models were reported in bold. Model 4, in both studies, did not converge. WLSMV = Weighted Least Squares Mean- and Variance-adjusted; *df* = degrees of freedom; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; RMSEA = Root Mean Square Error of Approximation; AIC = Akaike's Information Criterion; Δ AIC = difference in AIC (the best model is marked by 0); MFN = Method Factor associated with Negatively worded items

Results

Descriptive Statistics and Missing Data

The means for the 7-point Likert items ranged from 2.40 (Item 2, *God's love*) to 5.96 (Item 7, *family support*) with an overall mean of 4.39 ($SD = 0.91$). The distribution was fairly normal, indeed values for skewness ranged from -1.53 to 0.91 and values for kurtosis ranged from -0.94 to 2.90 . All items were strongly correlated with their respective scale total scores (r_{it} ; $M = .63$, $SD = .16$, minimum = $.30$, maximum = $.92$). Full details are presented in Table 1A (see Supplementary Material). For all items, the highest frequency of missingness was 2 and the minimum covariance coverage was 99.1%. Therefore, missingness was substantially negligible and therefore treated with pairwise method (as the default in *Mplus* when using categorical estimators).

First-Order Confirmatory Factor Analyses

Table 1 reports goodness-of-fit indices for alternative models. Model 5, including the seven correlated CSW factors and one method effect (see Fig. 1, Model 5) showed the best fit.

Completely standardized loadings for Model 5 are presented in Table 2. These loadings ranged from $.44$ to $.96$ ($M = .74$, $SD = .14$) for the seven CSW factors and from $.23$ to $.57$ ($M = .37$, $SD = .14$) for the method factor associated with negatively worded items (MFN). The Average Variance Extracted (AVE) was 0.15 for method factor, 0.88 for *God's love*, 0.46 for *virtue*, 0.55 for *academic competence*, 0.64 for *competition*, 0.47 for *approval from others*, 0.46 for *appearance*, 0.49 for *family support*. Correlations among latent factors (see the bottom part of Table 2) were all positive and significant at $p \leq .001$ except the correlations between *competition* with *God's love* ($r = -.07$, $z = -1.382$, $p = .167$), *appearance* with *God's love* ($r = -.06$, $z = -1.213$, $p = .225$), and *appearance* with *virtue* ($r = .06$, $z = 1.233$, $p = .218$). The remaining correlations ranged from $.17$ ($z = 3.462$, $p = .001$; *family support* with *appearance*) to $.65$ ($z = 23.272$, $p < .001$; *competition* with *academic competence*) with a mean of $.37$ ($SD = .15$).

Second-Order Confirmatory Factor Analyses

All our second-order confirmatory factor analyses started from the best fitting univariate solution (Model 5; seven

Table 2 Parameter Estimates from the Best Fitting First-Order Models (i.e., Model 5)

	Study 1		Study 2	
	λ	λ_{MFN}	λ	λ_{MFN}
<i>God's love</i>				
CSW2	.92	–	.94	–
CSW8	.95	–	.95	–
CSW18	.96	–	.96	–
CSW26	.95	–	.95	–
CSW31	.91	–	.90	–
<i>Virtue</i>				
CSW5	.51	–	.48	–
CSW11	.60	–	.63	–
CSW14	.69	–	.68	–
CSW28	.71	–	.71	–
CSW34	.84	–	.80	–
<i>Academic competence</i>				
CSW13	.61	.34	.67	.33
CSW19	.76	–	.71	–
CSW22	.82	–	.82	–
CSW27	.86	–	.87	–
CSW33	.63	–	.71	–
<i>Competition</i>				
CSW3	.80	–	.81	–
CSW12	.73	–	.77	–
CSW20	.84	–	.81	–
CSW25	.80	–	.84	–
CSW32	.82	–	.88	–
<i>Approval from others</i>				
CSW6	.69	.48	.65	.44
CSW9	.44	–	.35	–
CSW15	.69	.57	.62	.65
CSW23	.54	.46	.61	.41
CSW35	.97	–	.93	–
<i>Appearance</i>				
CSW1	.54	–	.71	–
CSW4	.65	.24	.71	.19
CSW17	.82	–	.77	–
CSW21	.62	–	.47	–
CSW30	.72	.24	.77	.15
<i>Family support</i>				
CSW7	.65	–	.67	–
CSW10	.46	.23	.51	.23
CSW16	.86	–	.83	–
CSW24	.73	–	.72	–
CSW29	.74	–	.74	–

Table 2 (continued)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
1. God's love	1	.33***	.18**	-.09 ^{n.s.}	.20***	-.06 ^{n.s.}	.28***
2. Virtue	.35***	1	.44***	.15**	.26***	.06 ^{n.s.}	.42***
3. Academic competence	.20***	.42***	1	.57***	.43***	.53***	.51***
4. Competition	-.07 ^{n.s.}	.21***	.65***	1	.35***	.61***	.25***
5. Approval from others	.18***	.23***	.49***	.47***	1	.50***	.34***
6. Appearance	-.06 ^{n.s.}	.06 ^{n.s.}	.50***	.58***	.52***	1	.19**
7. Family support	.31***	.40***	.49***	.24***	.27***	.17**	1

Note. All λs were significant at $p \leq .001$

λ = Standardized factor loading; MFN = Method Factor associated with Negatively worded items. Correlations below the diagonal refer to Study 1, correlations above the diagonal refer to Study 2

^{n.s.} = not statistically significant, or $p > .10$; ⁺ $p \leq .10$; * $p \leq .05$; ** $p \leq .01$; *** $p \leq .001$

specific factors and one uncorrelated method factor). Table 1 presents goodness of fit indices for the three second-order models (Models 6–8). The fit of Model 7 resulted in the best higher-order factor solution, since it fit the data better than the competing higher-order models (see Table 1). At the second-order level, loading values (see Fig. 2) ranged from .38 ($z = 7.241, p < .001$; *God's love* on the *private sphere* CSW factor) to .80 ($z = 26.325, p < .001$; *competition* on the *public sphere* CSW factor). The two higher-order factors of *private sphere* CSW and *public sphere* CSW were moderately correlated ($r = .33, z = 5.481, p < .001$).

Discussion

There were two important results offered by this study. First, items from the CSW scale were significantly contaminated by method variance. Accordingly, including a latent factor representing a method effect associated with negatively worded items is necessary not only for obtaining good-fitting models but also for obtaining a purer measure of the seven CSW domains. Second, the analysis of the second-order structure of the CSW scale revealed a significant deviation from the theoretical expectation that CSW domains should group into the two clusters initially hypothesized by Crocker et al. (2003). Indeed, we found evidence that the model empirically derived by Stefanone et al. (2011) best represented

the observed covariance among the seven first-order CSW domains.

One aspect of this modified model is particularly important. The dimension of *academic competence* was revealed to be a kind of middle ground posited at the intersection of private and public domains while being specific to neither of them. The uses of scores on the two higher-order dimensions could be challenging when one is interested in using the broad dimensions of *private sphere* and *public sphere* of CSW outside of a CFA model. In conclusion, we found evidence that a mixed second-order bifactor structure best represents the covariance between the CSWS items. In Study 2, we further tested the stability of this factor solution and evaluated the external correlates of each factor.

Study 2

The first objective of this study was to further evaluate the best fitting models from Study 1 (i.e., Model 5 and Model 7; see Fig. 1 and Fig. 1A) on a different sample of students. To this aim, we retested the entire sequence of models considered in Study 1. We formulated the following two hypotheses:

Hypothesis 1: Among the first-order competitive models, Model 5 is the best fitting one.

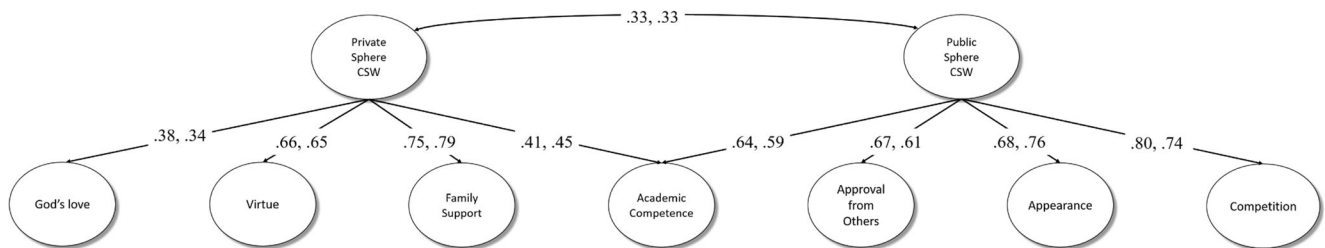


Fig. 2 The best fitting second-order model (Model 7, in Fig. 1A [in Supplementary Material] and in Table 1). Factor loadings are presented in standardized form and were all significant at $p < .001$. The first factor loading refers to Study 1, the second factor loading (after the comma)

refers to Study 2. Residual variances of first-order latent variables, measurement models, and method factor (MFN; see Model 5 in Fig. 1) were omitted for sake of clarity

Hypothesis 2: Among the second-order competitive models, Model 7 is the best fitting one.

The second and more important objective of this study was to investigate the external validity of the CSWS. To this aim, in a longitudinal design, we assessed CSW, personality traits, general self-esteem (GSE), implicit self-esteem (ISE), and religiosity in students before they began their sophomore year of college. Then we involved the same students in a daily study lasting 5 days in which they completed measures of daily self-esteem that we scored to obtain indices of self-esteem instability (GSE_{Inst}) and level (GSE_{Level}). At the end of the first semester, measures of grade point average (GPA) and depression were gathered for each student.

In accordance with correlations reported in previous studies (Crocker et al. 2003; Sargent et al. 2006; Stefanone et al. 2011; Zeigler-Hill et al. 2008), we predicted that *God's love* would correlate significantly and positively with religiosity (*Hypothesis 3*); *academic competence* would correlate significantly and positively with GPA (*Hypothesis 4*) and negatively with emotional stability (*Hypothesis 4a*); *virtue* would correlate significantly and positively with conscientiousness (*Hypothesis 5*) and agreeableness (*Hypothesis 5a*); *appearance* would correlate significantly and negatively with emotional stability (*Hypothesis 6*) and positively with depression (*Hypothesis 6a*); *family support* would correlate significantly and positively with agreeableness (*Hypothesis 7*); *competition* would correlate significantly and positively with GPA (*Hypothesis 8*); and *approval from others* would correlate significantly and negatively with GSE (*Hypothesis 9*).

At the second-order level, we hypothesized that *private sphere CSW* would correlate significantly and negatively with openness (*Hypothesis 10*) and positively with conscientiousness (*Hypothesis 10a*), whereas *public sphere CSW* would correlate significantly and positively with energy/extraversion (*Hypothesis 11*) and agreeableness (*Hypothesis 11a*). We did not specify any hypotheses regarding the relationships between CSW and GSE_{Inst} , GSE_{Level} , and ISE because few studies have addressed these relationships (e.g., Maroiu et al. 2016). Furthermore, we expected that the hypothesized correlations would range between a low (.10) and medium (.30) effect size (Cohen 1992) given that previous studies have usually shown moderately low correlations between CSW and external criteria (Crocker et al. 2003; Sargent et al. 2006; Stefanone et al. 2011; Zeigler-Hill et al. 2008). The only exception is the relationship between *God's love* and religiosity, for which we expect a high correlation ($> .50$; Cohen 1992).

Method

Participants and Procedure

This study was based on a convenience sample of 293 Italian undergraduates (77.1% females), all Caucasian, enrolled in two introductory psychology classes and compensated with partial course credit. The participation was voluntary and no exclusion criteria was selected. Their age ranged from 19 to 48 years with a mean of 20.99 ($SD = 3.38$). The participation was voluntary and no exclusion criteria were applied (nor any outlying participant was detected or excluded).

Participants' filled questionnaires on personality traits, religiosity, general self-esteem, and implicit self-esteem along with the CSWS administered by using an internal online platform. After two weeks, participants completed the modified version of the Rosenberg General Self-Esteem (RGSE) scale online at 24-h intervals (from 8 p.m. to 12 a.m.) for 5 consecutive days (e.g., from Monday to Friday). To enhance study participation, participants received an e-mail reminder at 7:55 p.m. each day with a link to a website to complete the daily RGSE scale. Approximately two months after the end of the daily surveys, participants received an additional booklet containing the Center for Epidemiologic Studies Depression (CES-D) scale and a question about their GPA.

Measures

Contingencies of Self-Worth Scale (CSWS). Participants filled out the same version of the CSWS as in Study 1. Reliability for each of the seven subscales was widely acceptable (all Cronbach's alphas and omegas were above .72; see Table 1A in Supplementary Material). **General Self-Esteem (GSE).** GSE was measured with the Rosenberg General Self-Esteem scale (Rosenberg 1965). Each item was scored on a 4-point Likert scale ranging from 1 (*Strongly disagree*) to 4 (*Strongly agree*). Cronbach's alpha reliability coefficient was .81. **Daily Self-Esteem.** Following the general procedure outlined by Kernis and his colleagues for measuring self-esteem instability (e.g., Kernis et al. 1993), participants were asked to complete a modified version of the RGSE scale each day. The RGSE scale was modified so that participants were instructed to give the response that best reflected how they felt at the moment that they completed the measure. Responses were offered on scales ranging from 1 (*Strongly disagree*) to 4 (*Strongly agree*). Reliability coefficients for days 1 through 5 were .86, .88, .89, .89, and .89, respectively. We computed the average self-esteem level (i.e., stable self-esteem; GSE_{Level}) as the mean of the daily score obtained by the participants on this scale and the self-esteem instability (GSE_{Inst}) as the standard deviation of each participant's score as observed across the 5-day period (see Kernis et al.

1993). **Implicit Self-Esteem (ISE).** The Implicit Association Test (IAT; Greenwald and Farnham 2000; Greenwald et al. 1998), implemented online through the INQUISIT software package (Millisecond Software 2000), was used to assess implicit self-esteem. In this test, the stimuli of the target-concept categories (Self vs. Others) were words related to “self” or “me” vs. “others” or “them”. The stimuli words for the attribute-dimension (Pleasant vs. Unpleasant) were the emotionally-loaded attributes (e.g., positive/good vs. negative/bad). In the IAT, the participants performed two types of categorization tasks with five stimuli for each category. The words were presented in random order within each block of trials. As described by Greenwald et al. (1998), the entire procedure consisted of seven blocks of trials. Blocks 1 (Self vs. Others), 2 (Pleasant vs. Unpleasant), and 5 (Others vs. Self) were single categorization blocks of 20 trials, whereas Blocks 3–4 and 6–7 were combined blocks (Self or Pleasant vs. Others or Unpleasant) of 20 (3–6) and 40 (4–7) trials. Participants were requested to respond as quickly and accurately as possible to the stimuli-words that appeared on the monitor. Following Greenwald et al. (2003), data from blocks 3–4 and 6–7 were used to compute IAT difference scores according to the built-in error penalty method. Positive scores indicated high implicit self-esteem, and negative scores indicated low implicit self-esteem. The internal consistency of the scale scores was .56. Reliability was estimated by a split-half index based on two partial scores, respectively computed from blocks 3 and 6 (20 + 20 trials) and from blocks 4 and 7 (40 + 40 trials) through the Spearman-Brown formula. **Personality Traits.** Personality traits were measured through a short version of the Big Five Questionnaire (BFQ; Caprara et al. 1993) containing 60 items (response scale ranged from 1 = *Very false for me* to 5 = *Very true for me*) that form five domain scales: energy/extraversion, friendliness/agreeableness, conscientiousness, emotional stability (vs. neuroticism), and openness. Reliability coefficients ranged from .73 (energy/extraversion) to .88 (emotional stability). **Depression.** Participants rated their levels of depressive symptoms using the CES-D, a 20-item scale developed by Radloff (1977). This scale measures the symptoms that characterize depression such as despondency, hopelessness, loss of appetite and interest in pleasurable activities, sleep disturbance, crying bouts, loss of initiative, and self-deprecation. The items were rated based on frequency of occurrence during the past week using a 4-point Likert-type scale, ranging from 1 = *Rarely or none of the time (less than 1 day)* to 4 = *Most or all of the time (5–7 days)*. An example item is: “*I was bothered by things that usually don’t bother me*”. The reliability coefficient was .88. **Religiosity.** Religiosity was measured with a single item (“*How religious are you?*”) rated on a scale ranging from 1 (“*Not at all*”) to 7 (“*Extremely*”).

Results

Descriptive Statistics and Missing Data

The means for the 7-point Likert items ranged from 2.44 (Item 2, *God’s love*) to 6.09 (Item 7, *family support*) with an overall mean of 4.45 ($SD = .93$). The distribution was fairly normal, indeed values for skewness ranged from -1.83 to 0.93 and values for kurtosis ranged from -1.08 to 4.92 . All items were strongly correlated with their respective scale total scores (r_{tt} : $M = .63$, $SD = .17$, minimum = $.23$, maximum = $.92$). Full details are presented in Table 1A (see Supplementary Material). For all items, the highest frequency of missingness was 2 and the minimum covariance coverage was 98.6%. As for Study 1, missingness was substantially negligible and therefore treated with pairwise method (as the default in *Mplus* when using categorical estimators).

First-Order and Second-Order Confirmatory Factor Analyses

The results of the model-fitting analyses fully replicated those obtained in Study 1 (see Table 1, Table 2, and Fig. 2). Model 5 (see Table 1) best fit the data at the first-order level, thus supporting *Hypothesis 1*. Factor loadings (see Table 2) ranged from .35 to .96 ($M = .74$, $SD = .15$) for the seven CSW factors and from .15 to .65 ($M = .34$, $SD = .17$) for the method factor associated with negatively worded items. The Average Variance Extracted (AVE) was 0.14 for method factor, 0.88 for *God’s love*, 0.45 for *virtue*, 0.58 for *academic competence*, 0.68 for *competition*, 0.44 for *approval from others*, 0.48 for *appearance*, 0.49 for *family support*. Correlations among latent factors (see the bottom part of Table 2) were all positive and significant at $p \leq .01$ except (consistent with Study 1) the correlations between *competition* with *God’s love* ($r = -.09$, $z = -1.573$, $p = .116$), *appearance* with *God’s love* ($r = -.06$, $z = -0.909$, $p = .363$), and *appearance* with *virtue* ($r = .06$, $z = 0.960$, $p = .337$). The remaining correlations ranged from .15 ($z = 2.616$, $p = .009$; *virtue* with *competition*) to .61 ($z = 15.000$, $p < .001$; *competition* with *appearance*) with a mean of .36 ($SD = .14$). At the higher-order level, Model 7 (see Table 1) provided the best data fit, thus supporting *Hypothesis 2*. Loading values (reported in Fig. 2) ranged from .34 ($z = 5.442$, $p < .001$; *God’s love* on the *private sphere* CSW factor) to .79 ($z = 11.445$, $p < .001$; *family support* on the *private sphere* CSW factor). The two higher-order factors of *private sphere* CSW and *public sphere* CSW were moderately correlated ($r = .33$, $z = 4.436$, $p < .001$).

Correlations with External Criteria

To investigate the construct validity of CSWS domains, individuals’ scores on the first-order CSW dimensions were

computed as the average of the items composing the dimension, whereas scores on the two higher-order dimensions were computed as the average of the first-order dimensions composing each second-order dimension after excluding *academic competence* from both scores, as suggested by Stefanone et al. (2011).

In general, the hypothesized correlations between measures of CSW with external criteria ranged from low to moderate, as hypothesized (see Table 3). As far as measures of self-esteem are concerned, implicit self-esteem (ISE) failed to show any significant correlation with any CSW dimensions, whereas general self-esteem (GSE), level of self-esteem (GSE_{Level}), and instability of self-esteem (GSE_{Inst}) showed a few significant correlations. Indeed, GSE correlated significantly and negatively only with *approval from others* ($r = -.13, p = .03$; thus confirming *Hypothesis 9*); GSE_{Level} was significantly and negatively correlated with *appearance* ($r = -.17, p = .004$) and *approval from others* ($r = -.19, p = .001$); and GSE_{Inst} was significantly and positively correlated with *competition* ($r = .16, p = .008$) and *appearance* ($r = .14, p = .016$). Depression (measured 2 months after the end of the study) showed two positive and significant correlations with *appearance* ($r = .19, p = .001$; thus confirming *Hypothesis 6a*) and *academic competence* ($r = .15, p = .014$). Religiosity showed four positive and significant correlations. In particular, it was strongly correlated with *God's love* ($r = .75, p < .001$; thus confirming *Hypothesis 3*) and *private sphere CSW* ($r = .63, p < .001$), and it was moderately correlated with *family love* ($r = .20, p = .001$) and *virtue* ($r = .20, p = .001$). Academic GPA (measured 2 months after the end of the study) showed three interesting positive and significant correlations with *academic competence* (thus confirming *Hypothesis 4*), *approval from others*, and *public sphere*

CSW (all $r_s = .14, p_s = .02$). Instead, contrary to our expectation, GPA did not significantly correlate with *competition* (so *Hypothesis 8* was not supported).

Among the Big Five traits, energy/extraversion, agreeableness, and conscientiousness showed several significant and positive correlations, whereas emotional stability and openness showed significant correlations with a negative sign. In particular, energy/extraversion showed positive and significant correlations with *competition* ($r = .24, p < .001$), *academic competence* ($r = .15, p = .009$), and *public sphere CSW* ($r = .13, p = .029$; thus supporting *Hypothesis 11*); agreeableness showed positive and significant correlations with *family support* ($r = .19, p = .001$; thus supporting *Hypothesis 7*), *God's love* ($r = .16, p = .005$), *virtue* ($r = .16, p = .007$; thus supporting *Hypothesis 5a*), and *private sphere CSW* ($r = .23, p < .001$; thus, our *Hypothesis 11a* regarding the positive relationship between agreeableness and *public sphere CSW* was not supported); conscientiousness showed positive and significant correlations with *God's love* ($r = .15, p = .011$), *virtue* ($r = .17, p = .004$; thus supporting *Hypothesis 5*), and *private sphere CSW* ($r = .19, p = .001$; thus supporting *Hypothesis 10a*); emotional stability showed negative and significant correlations with *family support* ($r = -.16, p = .008$), *appearance* ($r = -.24, p < .001$; thus supporting *Hypothesis 6*), *academic competence* ($r = -.23, p < .001$; thus supporting *Hypothesis 4a*), and *public sphere CSW* ($r = -.18, p = .003$); and openness to experience showed negative and significant correlations with *God's love* ($r = -.24, p < .001$) and *private sphere CSW* ($r = -.22, p < .001$; thus supporting *Hypothesis 10*). Confirming discriminant validity, each of the seven scales resulted significantly associated with their respective outcomes, and, in general, not with those outcomes related to the other subscales.

Table 3 Descriptive Statistics and Correlations Between CSW Scale Scores and External Criteria (Study 2)

	<i>M</i>	<i>SD</i>	GSE	GSE _{Level}	GSE _{Inst}	IAT	DEP	REL	GPA	E	A	C	S	O
1. Family support	5.34	0.90	.05	.08	-.08	.09	.01	.20**	-.01	.03	.19**	.09	-.16**	-.09
2. Competition	4.73	1.15	-.01	.03	.16**	-.09	.04	-.11	.03	.24***	-.09	-.11	-.10	-.09
3. Appearance	4.66	1.05	-.07	-.17**	.14*	-.01	.19**	-.09	.09	.04	-.11	-.11	-.24***	-.03
4. God's love	2.72	1.67	-.05	-.02	-.06	.03	-.05	.75***	.08	-.01	.16**	.15*	.07	-.24***
5. Academic competence	4.99	0.97	.01	-.01	.03	.04	.15*	.07	.14*	.15**	.08	.10	-.23***	-.06
6. Virtue	4.76	0.94	.03	-.01	0	0	.01	.20**	.08	.02	.16**	.17**	.03	-.10
7. Approval from others	3.99	1.12	-.13*	-.19**	-.07	.02	.06	.07	.14*	-.09	.02	.03	-.08	-.07
8. Private sphere CSW	4.28	0.85	0	.01	-.07	.05	-.03	.63***	.08	.02	.23***	.19**	0	-.22***
9. Public sphere CSW	4.57	0.81	-.06	-.08	.05	-.02	.10	.01	.14*	.13*	0	.01	-.18**	-.09

Note. *M* = Mean; *SD* = Standard deviation; GSE = General Self-Esteem; GSE_{Level} = Observed average level of self-esteem across the five days; GSE_{Inst} = Observed within-individual standard deviation of self-esteem scores across the five days (state self-esteem); IAT = Implicit self-esteem; DEP = Depression; REL = Religiosity; GPA = Grade Point Average; E = Energy/Extraversion; A = Agreeableness; C = Conscientiousness; S = Emotional Stability; O = Openness to experiences; CSW = Contingencies of Self-Worth

* $p < .05$; ** $p < .01$; *** $p < .001$. Correlations without apex were nonsignificant ($p > .05$). All correlations statistically significant at least at $p < .05$ were bolded

Ancillary Analyses

Measurement Invariance across Studies We conducted a measurement invariance analysis between samples from Study 1 and Study 2 for the second-order best fitting model (i.e., Model 7). First of all, we checked the presence of differences between the number of categories for each item across samples. We found that item 7 and item 16 had a different number of categories; indeed, we found an empty cell for “item 7 category 3” and for “item 16 category 1” in Study 2. Given that the frequency for those cells was negligible in Study 1 (3 and 1, respectively) we collapsed those categories with the subsequent ones (i.e., we recoded category 3 with 4 for item 7 and category 1 with 2 for item 16). Then, we conducted a multiple-group measurement invariance analysis for ordinal data (Bowen and Masa 2015). We started by testing configural invariance simply assuming that Model 7 fits well in both samples. Indices of fit supported this first step (WLSMV-Based $\chi^2 = 2727.594$, $df = 1088$, $p < .001$; CFI = .966; TLI = 0.963; RMSEA = 0.064). Then, in order to test metric invariance, we constrained first-order factor loadings (i.e., factor loadings linking observed variables to the seven specific factors and method factor) to be equal across samples. The *diff-test function* (implemented in *Mplus* for comparing models estimated with WLSMV) showed a significant worsening of the fit ($\Delta\chi^2 = 52.978$, $df = 34$, $p = .0201$). Indeed, we found that the factor loading linking Appearance to item 21 was significantly different across samples. After removing that constraint, the model showed a good fit (WLSMV-Based $\chi^2 = 2680.341$, $df = 1121$, $p < .001$; CFI = .967; TLI = 0.965; RMSEA = 0.061) and a non-significant *diff-test* ($\Delta\chi^2 = 42.503$, $df = 33$, $p = .1243$). Then, given the second-order structure of Model 7 (Chen et al. 2005), we tested the invariance of the factor loading linking the higher-order factors to the lower-order factors. Indices of fit (WLSMV-Based $\chi^2 = 2509.310$, $df = 1127$, $p < .001$; CFI = .971; TLI = 0.969; RMSEA = 0.057) and *diff-test* ($\Delta\chi^2 = 6.070$, $df = 6$, $p = .4154$) supported metric invariance of second-order factor loadings. Finally, we tested strong invariance by constraining all thresholds to be equal across samples. Again, indices of fit (WLSMV-Based $\chi^2 = 2602.568$, $df = 1334$, $p < .001$; CFI = .973; TLI = 0.976; RMSEA = 0.050) and *diff-test* ($\Delta\chi^2 = 193.070$, $df = 207$, $p = .7478$) supported this last step of invariance analysis.

Exploratory Structural Equation Model As suggested by an anonymous Reviewer, we conducted an Exploratory Structural Equation Model (ESEM) to the best fitting first-order solution (i.e., Model 5), in order to strengthen the validity of this model. Indeed, ESEM allows to estimate all the cross-loadings that in a common CFA are constrained to be zero (e.g., Morin and Maïano 2011). Hence, ESEM allows a thorough investigation of the relationship between all latent

factors and all observed indicators, without zero-loading constraints. Thus, we estimated an ESEM starting from Model 5 structure, by specifying a “target” rotation, so that cross-loadings are estimated as close to zero as possible. Results indicate a good fit to data. For Study 1: WLSMV-based $\chi^2(364) = 855.438$, $p < .001$; CFI = 0.983; TLI = 0.972; RMSEA = 0.055 (90%CI = 0.050, 0.059). For Study 2: WLSMV-based $\chi^2(364) = 644.656$, $p < .001$; CFI = 0.985; TLI = 0.976; RMSEA = 0.051 (90%CI = 0.045, 0.058). However, the most interesting information are provided by inspection of factor loadings values. As it can be seen in Table 2A (Supplementary Material) target loading were all satisfactory whereas cross-loading are all low in magnitude. Thus, the ESEM solution further supports the validity of Model 5 structure.

Discussion

In Study 2, we confirmed results from Study 1 suggesting that the model with one method factor along with the seven substantive factors best fit the data. Moreover, we found further evidence that, at the second-order level, the covariance among the seven dimensions was best represented by two general factors grouping them in *public sphere* and *private sphere* of CSW.

An important aim of this study was to investigate the external validity of the seven CSW factors and the two higher-order factors. Our results showed a sufficient degree of external validity for each dimension included in the model. Indeed, 12 out of 14 hypotheses regarding the correlations between CSW domains and external criteria were confirmed in both their direction and size. Furthermore, the explorative analysis on the relationship between CSW and GSE_{Level} , GSE_{Inst} , and ISE highlighted the orthogonal relationship between CSW and these other types of self-esteem, which should be further investigated in future research.

General Discussion

We conducted two studies to provide a psychometric analysis regarding the internal and external validity of the first- and second-order structure of the Contingencies of Self-Worth Scale (CSWS). To this aim, we submitted the instrument to an in-depth psychometric investigation, comparing alternative measurement models and using several external criteria to assess the external validity of scale scores. Overall, our results suggest the tenability of the seven-factor structure when one considers an additional orthogonal method factor capturing spurious variance introduced by the presence of negatively worded items. Ignoring this additional method factor results in a model with a sub-optimal fit and biased factor loading estimates.

While confirming the first-order structure of the scale, our study also suggested that, at the second-order level, the relationships among the seven factors might differ from what was originally proposed by Crocker et al. (2003). Across two studies, the original model, proposing a dichotomy between internal/intrapersonal (i.e., *God's love* and *virtue*) and external/interpersonal (i.e., *academic competence*, *competition*, *approval from others*, *appearance*, and *family support*) contingencies, fared worse in terms of data fit than the revised model proposed by Stefanone et al. (2011), which contrasted *private sphere* (i.e., *family support*, *virtue*, *God's love*, and *academic competence*) and *public sphere* (i.e., *approval from others*, *appearance*, *competition*, and *academic competence*) of CSW.

This result is not devoid of complications given that, under this model, the *academic competence* domain of CSW is placed at the intersection between *private sphere* and *public sphere* of CSW. From a conceptual point of view, it is conceivable that the academic context represents an intermediate domain reflecting partly what the individual feels or strives to be (i.e., a good student in order to get a rewarding job) and partly what the individual in that specific moment is for society (i.e., a university student). In any case, this seems to be an area in need of further conceptual and theoretical developments from future research.

From a practical point of view, the cross-loading items of this dimension pose challenges to applied researchers. This is likely an area in need of further research and conceptual development before formulating any definitive recommendations. In the present study, we computed the individuals' scores on the *private sphere CSW* and the *public sphere CSW* after excluding *academic competence* from both dimensions. While this seems like a recommended (see Stefanone et al. 2011) and practical solution, it is not psychometrically sound. As a matter of fact, this procedure introduces a discrepancy between the structural higher-order model (in which the dimension exists) and the observed score constructed (in which the dimension is ignored). In sum, future studies should go further in proposing revisions of the CSWS aimed to clarify the nature of the *academic competence* dimension as belonging to the *public sphere* or the *private sphere* of CSW.

Finally, our Study 2 provided support for the external validity of the seven first-order dimensions and the two higher-order dimensions of CSWS. Surprisingly, we found few significant correlations between all the CSW dimensions and different measures of self-esteem. This might suggest that CSW are substantially different from general self-esteem (GSE) because, whereas GSE has mostly trait-like characteristics, CSW is conceptualized to be more prone to change according to variation in one's capability of self-regulation (e.g., Crocker et al. 2006). Future studies should investigate and explain these differences (e.g., through genetic studies).

In general, as previously said, most of our hypothesized correlations were in the suggested direction, therefore supporting a certain degree of external validity; however, the low-size correlations between some CSWS domains and their hypothesized external criteria is worth of note. Indeed, although this result mimics patterns in previous works (Crocker et al. 2003), one may speculate about the potential causes of this problem. For example, one may wonder if it reflects a theoretical misalignment between the theoretical status of the CSW and the outcomes considered (that reflects those used in the seminal work by Crocker et al. 2003), or if it suggests the necessity of a more deepen analysis of the CSW scale items content, or, simply, if it points to the need for a theoretically oriented reconsideration of CSW correlates. In a similar vein, recently Briganti et al. (2019) conducted a network analysis on the CSWS and find that "the seven domains of self-worth form a heterogeneous system in which domains are not uniformly positively connected with each other" (Briganti et al. 2019, p. 255). Considering together Briganti et al.'s (2019) findings and our main results (an ambiguous second-order bifactor structure and the small correlations with external variables), it seems reasonable to ponder if the CSWS dimensions are measures of self-worth. However, given that this is outside the scope of the present contribution, we recommend that this problem becomes the aim of future and dedicated studies on the content validity of the CSWS items. Indeed, the latter is a pivotal point to be addressed in order to correctly operationalize the construct of contingencies of self-worth. Another important point regards the need of managing the presence of negatively worded items. Given that there are no clear procedure to take care of these effects using observed scores, it is likely that observed scores obtained from the CSWS scale might be contaminated. Finally, whereas our two-study contribution aimed at testing the most commonly used competing ways to modeling the internal structure of the CSWS (hence, we simply looked for the best solution among those selected models), we point out that best-fitting models (i.e., Model 5 and Model 7) are not without problems. For instance, all WLSMV-based χ^2 were significant; while it is a common procedure to rely on approximate fit indices more than on χ^2 test statistic (Ropovik 2015), a significant value of the latter indicates the presence of some model misspecifications. Our inspection of residuals and modification indices (Ropovik 2015) showed that some residual covariances among items should be freed, as well as some cross-loadings. In conclusion, while we found that Model 5 and 7 represent the best solution among inspected models, we do not underestimate the possibility that future new dimensional conceptualizations and/or new reevaluations of item content are needed, in order to improve the structural validity of the CSWS. However, we believe that our contribution is a good starting point for future refinements of the scale.

Putting all said together, our findings and those of previous research may provide useful information for a future reconsideration of the structure and the item content of an instrument aiming at correctly assessing contingencies of self-worth.

Limitations

Several limitations should be acknowledged. First, some constraints on the generalizability of our findings deserve to be discussed (Simons et al. 2017). Indeed, the samples we used were recruited from the university population, so it is not possible to generalize our results to other populations. Yet, future studies should investigate whether our findings could be replicated in clinical samples or in samples with older people. Finally, we conducted both studies in Italy; hence, our findings could be generalizable at the best to other similar western cultures, but given the cultural differences in terms of contingencies of self-worth, our findings cannot be generalized to other different cultures, such as eastern countries.

Second, the CSWS was analyzed at a cross-sectional level. Future studies should test the longitudinal tenability of the CSWS structure found in our studies in terms of longitudinal measurement invariance.

Third, in this study we did not provide evidence concerning a possible substantial validity of the method factors. This might be an important task for future studies given that the literature offers similar examples (e.g., Alessandri et al. 2010, 2011).

Conclusion

The CSWS is a widely used self-report measure assessing self-worth across different domains and life situations. In this contribution, through two studies using large samples of university students, we attested (a) the validity of the seven first-order domains, (b) the validity of two higher-order spheres of CSW (*public* and *private*) that significantly accounted for the shared variance of the first-order domains, (c) the importance of including one further orthogonal factor composed of negatively worded items, and (d) the external validity of all dimensions. In this way, we hope that our study might advance the literature on CSWS and stimulate future research on this instrument, which constitutes an important tool in personality assessment.

Funding Open access funding provided by Università degli Studi di Roma La Sapienza within the CRUI-CARE Agreement. Enrico Perinelli was supported by a research fellowship (Research title: “WeBeWo LAB: Quality of Social Relationships in Workplace”, Number: 40103205) granted by the Autonomous Province of Trento and the Department of Psychology and Cognitive Science (University of Trento).

Guido Alessandri Was Supported by Two Grants from Sapienza University of Rome (RM11715C809391B1, RG11816433CBD8D3) Named “Progetti di Ateneo” (2017, 2018)

Compliance with Ethical Standards

Conflict of Interest All authors declares that they have no conflict of interest.

Ethical Approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed Consent Informed consent was obtained from all individual participants included in the study.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Alessandri, G., Vecchione, M., Eisenberg, N., & Laguna, M. (2015). On the factor structure of the Rosenberg (1965) general self-esteem scale. *Psychological Assessment*, 27(2), 621–635. <https://doi.org/10.1037/pas0000073>.
- Alessandri, G., Vecchione, M., Fagnani, C., Bentler, P., Barbaranelli, C., Medda, E., Nisticò, L., Stazi, M. A., & Caprara, G. V. (2010). Much more than model fitting? Evidence for the heritability of method effect associated with positively worded items of the life orientation test revised. *Structural Equation Modeling: A Multidisciplinary Journal*, 17(4), 642–653. <https://doi.org/10.1080/10705511.2010.510064>.
- Alessandri, G., Vecchione, M., Tisak, J., & Barbaranelli, C. (2011). Investigating the nature of method factors through multiple informants: Evidence for a specific factor? *Multivariate Behavioral Research*, 46(4), 625–642. <https://doi.org/10.1080/00273171.2011.589272>.
- Banks, H. T., & Joyner, M. L. (2017). AIC under the framework of least squares estimation. *Applied Mathematics Letters*, 74, 33–45. <https://doi.org/10.1016/j.aml.2017.05.005>.
- Baumgartner, H., & Steenkamp, J. B. E. (2001). Response styles in marketing research: A cross-national investigation. *Journal of Marketing Research*, 38(2), 143–156. <https://doi.org/10.1509/jmkr.38.2.143.18840>.
- Benson, P., & Spilka, B. (1973). God image as a function of self-esteem and locus of control. *Journal for the Scientific Study of Religion*, 12(3), 297–310. <https://doi.org/10.2307/1384430>.
- Bentea, C. C. (2016). Psychometrical properties of the contingencies of self-worth scale on a Romanian student sample. *Romanian Journal of Experimental Applied Psychology*, 7(1), 340–343.

- Blaine, B., & Crocker, J. (1995). Religiousness, race, and psychological well-being: Exploring social psychological mediators. *Personality and Social Psychology Bulletin*, 21(10), 1031–1041. <https://doi.org/10.1177/01461672952110004>.
- Bowen, N. K., & Masa, R. D. (2015). Conducting measurement invariance tests with ordinal data: A guide for social work researchers. *Journal of the Society for Social Work and Research*, 6(2), 229–249. <https://doi.org/10.1086/681607>.
- Briganti, G., Fried, E. I., & Linkowski, P. (2019). Network analysis of contingencies of self-worth scale in 680 university students. *Psychiatry Research*, 272, 252–257. <https://doi.org/10.1016/j.psychres.2018.12.080>.
- Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). New York: Guilford.
- Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: Understanding AIC and BIC in model selection. *Sociological Methods & Research*, 33(2), 261–304. <https://doi.org/10.1177/0049124104268644>.
- Caprara, G. V., Barbaranelli, C., Borgogni, L., & Perugini, M. (1993). The “big five questionnaire”: A new questionnaire to assess the five factor model. *Personality and Individual Differences*, 15(3), 281–288. [https://doi.org/10.1016/0191-8869\(93\)90218-R](https://doi.org/10.1016/0191-8869(93)90218-R).
- Chen, F. F., Sousa, K. H., & West, S. G. (2005). Teacher’s corner: Testing measurement invariance of second-order factor models. *Structural Equation Modeling: A Multidisciplinary Journal*, 12(3), 471–492. https://doi.org/10.1207/s15328007sem1203_7.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155–159. <https://doi.org/10.1037/0033-2909.112.1.155>.
- Cooley, C. H. (1902). *Human nature and the social order*. New York: Scribner’s.
- Coopersmith, S. (1967). *The antecedents of self-esteem*. San Francisco: Freeman.
- Crocker, J., Brook, A. T., Niiya, Y., & Villacorta, M. (2006). The pursuit of self-esteem: Contingencies of self-worth and self-regulation. *Journal of Personality*, 74(6), 1749–1772. <https://doi.org/10.1111/j.1467-6494.2006.00427.x>.
- Crocker, J., & Luhtanen, R. K. (2003). Level of self-esteem and contingencies of self-worth: Unique effects on academic, social, and financial problems in college students. *Personality and Social Psychology Bulletin*, 29(6), 701–712. <https://doi.org/10.1177/0146167203029006003>.
- Crocker, J., Luhtanen, R. K., Cooper, M. L., & Bouvrette, A. (2003). Contingencies of self-worth in college students: Theory and measurement. *Journal of Personality and Social Psychology*, 85(5), 894–908. <https://doi.org/10.1037/0022-3514.85.5.894>.
- Crocker, J., & Wolfe, C. T. (2001). Contingencies of self-worth. *Psychological Review*, 108(3), 593–623. <https://doi.org/10.1037/0033-295X.108.3.593>.
- Cross, S., & Madson, L. (1997). Models of the self: Self-construals and gender. *Psychological Bulletin*, 122(1), 5–37. <https://doi.org/10.1037/0033-2909.122.1.5>.
- Demo, D. H., & Parker, K. D. (1987). Academic achievement and self-esteem among black and white college students. *Journal of Social Psychology*, 127(4), 345–355. <https://doi.org/10.1080/00224545.1987.9713714>.
- DiStefano, C., Liu, J., Jiang, N., & Shi, D. (2018). Examination of the weighted root mean square residual: Evidence for trustworthiness? *Structural Equation Modeling: A Multidisciplinary Journal*, 25(3), 453–466. <https://doi.org/10.1080/10705511.2017.1390394>.
- Eid, M., Lischetzke, T., Nussbeck, F. W., & Trierweiler, L. I. (2003). Separating trait effects from trait-specific method effects in multitrait-multimethod models: A multiple-indicator CT-C (M-1) model. *Psychological Methods*, 8(1), 38–60. <https://doi.org/10.1037/1082-989X.8.1.38>.
- Finney, S. J., & DiStefano, C. (2013). Nonnormal and categorical data in structural equation modeling. In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: A second course* (2nd ed., pp. 439–492). Charlotte, NC: Information Age Publishing.
- Geiser, C., Eid, M., West, S. G., Lischetzke, T., & Nussbeck, F. W. (2012). A comparison of method effects in two confirmatory factor models for structurally different methods. *Structural Equation Modeling: A Multidisciplinary Journal*, 19(3), 409–436. <https://doi.org/10.1080/10705511.2012.687658>.
- Geiser, C., & Lockhart, G. (2012). A comparison of four approaches to account for method effects in latent state–trait analyses. *Psychological Methods*, 17(2), 255–283. <https://doi.org/10.1037/a0026977>.
- Greenwald, A. G., & Farnham, S. D. (2000). Using the implicit association test to measure self-esteem and self-concept. *Journal of Personality and Social Psychology*, 79(6), 1022–1038. <https://doi.org/10.1037/0022-3514.79.6.1022>.
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, 74(6), 1464–1480. <https://doi.org/10.1037/0022-3514.74.6.1464>.
- Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the implicit association test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology*, 85(2), 197–216. <https://doi.org/10.1037/0022-3514.85.2.197>.
- Harter, S. (1986). Processes underlying the construction, maintenance, and enhancement of the self-concept in children. In J. Suls & A. W. Greenwald (Eds.), *Psychological perspectives on the self* (Vol. 3, pp. 136–182). Hillsdale, NJ: Erlbaum.
- Hintz, F., Geiser, C., & Shiffman, S. (2019). A latent state–trait model for analyzing states, traits, situations, method effects, and their interactions. *Journal of Personality*, 87(3), 434–454. <https://doi.org/10.1111/jopy.12400>.
- Hoge, D. R., Smit, E. K., & Hanson, S. L. (1990). School experiences predicting changes in self-esteem of sixth- and seventh-grade students. *Journal of Educational Psychology*, 82(1), 117–127. <https://doi.org/10.1037/0022-0663.82.1.117>.
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1–55. <https://doi.org/10.1080/10705519909540118>.
- James, W. (1890). *The principles of psychology* (Vol. 1). Cambridge, MA: Harvard University Press.
- Josephs, R. A., Markus, H. R., & Tafarodi, R. W. (1992). Gender and self-esteem. *Journal of Personality and Social Psychology*, 63(3), 391–402. <https://doi.org/10.1037/0022-3514.63.3.391>.
- Kempenaers, C., De Boeck, P., Dehon, C., Braun, S., & Linkowski, P. (2014). Metric invariance of the French version of the contingencies of self-worth scale. *Comprehensive Psychology*, 3, 1–13. <https://doi.org/10.2466/08.CP.3.22>.
- Kernis, M. H., Cornell, D. P., Sun, C.-R., Berry, A., & Harlow, T. (1993). There’s more to self-esteem than whether it is high or low: The importance of stability of self-esteem. *Journal of Personality and Social Psychology*, 65(6), 1190–1204. <https://doi.org/10.1037/0022-3514.65.6.1190>.
- Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed.). New York, NY: The Guilford Press.
- Isiordia, M., Conger, R., Robins, R. W., & Ferrer, E. (2017). Using the factor of curves model to evaluate associations among multiple family constructs over time. *Journal of Family Psychology*, 31(8), 1017–1028. <https://doi.org/10.1037/fam0000379>.
- Luhtanen, R. K., & Crocker, J. (2005). Alcohol use in college students: Effects of level of self-esteem, narcissism, and contingencies of self-worth. *Psychology of Addictive Behaviors*, 19(1), 99–103. <https://doi.org/10.1037/0893-164X.19.1.99>.
- Maricuțoiu, L. P., Macinga, I., Rusu, S., Virgă, D., & Sava, F. A. (2012). Adaptation and validation of the contingencies of self-worth scale

- on a Romanian student sample. *Cognition, Brain, Behavior. An Interdisciplinary Journal*, 16(1), 121–138.
- Maroïu, C., Maricuțoiu, L. P., & Sava, F. A. (2016). Explicit self-esteem and contingencies of self-worth: The moderating role of implicit self-esteem. *Personality and Individual Differences*, 99, 235–241. <https://doi.org/10.1016/j.paid.2016.05.022>.
- Mead, G. H. (1934). *Mind, self, and society*. Chicago: University of Chicago Press.
- Software, M. (2000). *Inquisit [computer software]*. Seattle, WA: Author.
- Morin, A. J., & Maïano, C. (2011). Cross-validation of the short form of the physical self-inventory (PSI-S) using exploratory structural equation modeling (ESEM). *Psychology of Sport and Exercise*, 12(5), 540–554. <https://doi.org/10.1016/j.psychsport.2011.04.003>.
- Muthén, L. K., & Muthén, B. O. (1998–2017). *Mplus User's Guide* (Eighth ed.). Los Angeles, CA: Author.
- Park, L. E., Crocker, J., & Kiefer, A. K. (2007). Contingencies of self-worth, academic failure, and goal pursuit. *Personality and Social Psychology Bulletin*, 33(11), 1503–1517. <https://doi.org/10.1177/0146167207305538>.
- Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of social psychological attitudes, Measures of personality and social psychological attitudes* (Vol. 1, pp. 17–59). San Diego, CA, US: Academic Press. <https://doi.org/10.1016/B978-0-12-590241-0.50006-X>.
- Podsakoff, P. M., MacKenzie, S. B., & Podsakoff, N. P. (2012). Sources of method bias in social science research and recommendations on how to control it. *Annual Review of Psychology*, 63, 539–569. <https://doi.org/10.1146/annurev-psych-120710-100452>.
- Radloff, L. S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement*, 1(3), 385–401. <https://doi.org/10.1177/014662167700100306>.
- Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research*, 47(5), 667–696. <https://doi.org/10.1080/00273171.2012.715555>.
- Richman, C. L., Brown, K. P., & Clark, M. (1987). Personality changes as a function of minimum competency test success or failure. *Contemporary Educational Psychology*, 12(1), 7–16. [https://doi.org/10.1016/S0361-476X\(87\)80034-6](https://doi.org/10.1016/S0361-476X(87)80034-6).
- Rindskopf, D., & Rose, T. (1988). Some theory and applications of confirmatory second-order factor analysis. *Multivariate Behavioral Research*, 23(1), 51–67. https://doi.org/10.1207/s15327906mbr2301_3.
- Ropovik, I. (2015). A cautionary note on testing latent variable models. *Frontiers in Psychology*, 6, 1715. <https://doi.org/10.3389/fpsyg.2015.01715>.
- Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: University Press. <https://doi.org/10.1515/9781400876136>.
- Rosenberg, M., Schooler, C., Schoenbach, C., & Rosenberg, F. (1995). Global self-esteem and specific self-esteem: Different concepts, different outcomes. *American Sociological Review*, 60(1), 141–156. <https://doi.org/10.2307/2096350>.
- Saleh, S. (2014). Robust AIC with high breakdown scale estimate. *Journal of Applied Mathematics*, 2014, 1–7. <https://doi.org/10.1155/2014/286414>.
- Sargent, J. T., Crocker, J., & Luhtanen, R. K. (2006). Contingencies of self-worth and depressive symptoms in college students. *Journal of Social and Clinical Psychology*, 25(6), 628–646. <https://doi.org/10.1521/jscp.2006.25.6.628>.
- Self and Social Motivation Lab. (2015). Contingencies of self-worth (CSW) scale. Found at <http://faculty.psy.ohio-state.edu/crocker/lab/csw.php#13> on May 26, 2018.
- Simons, D. J., Shoda, Y., & Lindsay, D. S. (2017). Constraints on generality (COG): A proposed addition to all empirical papers. *Perspectives on Psychological Science*, 12(6), 1123–1128. <https://doi.org/10.1177/1745691617708630>.
- Shrauger, J. S., & Schoeneman, T. J. (1979). Symbolic interactionist view of self-concept: Through the looking glass darkly. *Psychological Bulletin*, 86(3), 549–573. <https://doi.org/10.1037/0033-2909.86.3.549>.
- Solomon, S., Greenberg, J., & Pyszczynski, T. (1991). A terror management theory of social behavior: The psychological functions of self-esteem and cultural worldviews. *Advances in Experimental Social Psychology*, 24, 93–159. [https://doi.org/10.1016/S0065-2601\(08\)60328-7](https://doi.org/10.1016/S0065-2601(08)60328-7).
- Spilka, B., Shaver, P., & Kirkpatrick, L. A. (1985). A general attribution theory for the psychology of religion. *Journal for the Scientific Study of Religion*, 24(1), 1–20. <https://doi.org/10.2307/1386272>.
- Stefanone, M. A., Lackaff, D., & Rosen, D. (2011). Contingencies of self-worth and social-networking-site behavior. *Cyberpsychology, Behavior and Social Networking*, 14(1–2), 41–49. <https://doi.org/10.1089/cyber.2010.0049>.
- Yamaoka, K., Nakagawa, T., & Uno, T. (1978). Application of Akaike's information criterion (AIC) in the evaluation of linear pharmacokinetic equations. *Journal of Pharmacokinetics and Biopharmaceutics*, 6(2), 165–175. <https://doi.org/10.1007/BF01117450>.
- Wylie, R. (1979). *The self-concept: Theory and research on selected topics* (Vol. 2, 2nd ed.). Lincoln: University of Nebraska Press.
- Zeigler-Hill, V., Clark, C. B., & Pickard, J. D. (2008). Narcissistic subtypes and contingent self-esteem: Do all narcissists base their self-esteem on the same domains? *Journal of Personality*, 76(4), 753–774. <https://doi.org/10.1111/j.1467-6494.2008.00503.x>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.