Clarifying the Factor Structure of the Self-Compassion Scale: Nested Comparisons of Six Confirmatory Factor Analysis Models

Authors:

Noelle J. Strickland (Dalhousie University*),

Raquel Nogueira-Arjona (Dalhousie University*),

Sean Mackinnon (Dalhousie University*),

Christine Wekerle (McMaster University. 1200 Main St W, Hamilton, ON L8N 3Z5), &

Sherry H. Stewart (Dalhousie University*)

Notice: This is an electronic post-print of an article published in the European Journal of Psychological Assessment. This article may not exactly replicate the final version published in the journal but does reflect the article's content after the final peer review.

Corresponding author: Noelle Strickland (Department of Psychology and Neuroscience). Email: Noelle.strickland@dal.ca. *Mailing address: Dalhousie University Life Sciences Centre 1355 Oxford Street PO Box 15000 Halifax, NS B3H 4R2.

Abstract

Self-compassion is associated with greater well-being and lower psychopathology. There are mixed findings regarding the factor structure and scoring of the Self-Compassion Scale (SCS). Using confirmatory factor analysis, we tested and conducted nested comparisons of six previously posited factor structures of the SCS. Participants were *N*=1158 Canadian undergraduates (72.8% women, 26.6% men, 0.6% non-binary; M_{age}=19.0 [*SD*=2.3]). Results best supported a two-factor hierarchical model with six lower-order factors. A general self-compassion factor was not supported at the higher- or lower-order levels; thus, a single total score is not recommended. Given the hierarchical structure, researchers are encouraged to use structural equation models of the SCS with two latent variables: self-caring and self-coldness. A strength of this study is the large sample, while the undergraduate sample may limit generalizability.

Keywords: self-compassion, confirmatory factor analysis, Self-Compassion Scale, nested comparison, hierarchical factor structure.

Clarifying the Factor Structure of the Self-Compassion Scale: Nested Comparisons of Six Confirmatory Factor Analysis Models

Self-compassion – treating oneself with the same kindness as one would others – is a protective factor following trauma and a promising target in psychological therapy (Zeller et al., 2015). However, disagreement about the factor structure and validity of total self-compassion scores of the Self-Compassion Scale (SCS; Neff, 2003b), a popular self-compassion measure, warrants clarification, the objective of the current study.

Neff (2003a) posits a nested structure with six subdimensions (self-kindness, mindfulness, common humanity, self-judgement, overidentification, and isolation), three subdimensions, each with two corresponding positive and negative components (self-kindness/self-judgement, mindfulness/overidentification, common humanity/isolation), and one over-arching self-compassion factor, suggesting a one-, three-, or six-factor lower-order model may represent the construct. Gilbert (2010) argues self-compassion is comprised of factors aligned with the safeness processing system (self-caring; positive items) and the threat/defense processing system (self-coldness; negative items), suggesting a two-factor lower-order model.

In studies examining theoretically plausible factor structures of the SCS at the lower-order level, neither one- nor three-factor models were supported (e.g., Brenner et al., 2017; Coroiu et al., 2018; Petrocchi et al., 2014). A six-factor structure was supported in several studies (*N*=424, Williams et al., 2014; *N*=1554, Kotsou & Leys, 2016), but others report substantial model misfit (Coroiu et al., 2018; Costa et al., 2016; Zeng et al., 2016). After modifications, Costa et al. (2016; *N*=1263) supported two factors (self-caring and self-coldness), as did Lopez et al. (2015; *N*=1643) using exploratory factor analysis.

Supporting hierarchical models, Neff (2003b; N_1=391, N_2=232) reported six lower-order factors loading on a single higher-order self-compassion factor, though subsequent studies failed to replicate this structure (Castilho et al., 2015; Williams et al., 2014). Conversely, others found that two higher-order factors (self-caring, self-coldness) with six lower-order factors showed marginal (*N*=2448, Coroiu et al., 2018) to good fit (*N*=1115, Brenner et al., 2017; N_1=406, N_2=416, Montero-Marín et al., 2016).

Recently, the SCS factor structure has been examined using bifactor models, where all factors are estimated as direct pathways to items (Markon, 2019). Bifactor models with two general factors (self-caring, self-coldness) and six specific factors (Brenner et al., 2017; Coroiu et al., 2018) and with only one general factor and six specific factors were supported (Cleare et al., 2018; *N*=526; Neff et al., 2017; four samples; Tóth-Király et al., 2017; *N*=505). There are numerous reasons why we do not use bifactor models in the present study. In simulations, Murray and Johnson (2013) found that the bifactor model has more favorable fit indices than any other model, even when the true model is hierarchical. Bonifay and Cai (2017) found that bifactor models typically overfit the data and tend to have favorable fit indices when modelling random data. Reise et al. (2016) similarly caution readers that bifactor models can accommodate implausible and invalid data patterns (see also Markon, 2019). Moreover, Neff (2003b) conceptualized the higher-order SCS factors are explained by, or meaningful in the context of their association with, the lower-order factors. Given bifactor general and specific factors are orthogonal, clinically relevant factor interpretation is difficult (e.g., the meaning of self-caring exclusive of self-kindness; Bonifay et al., 2016). Therefore, bifactor analyses should be cautiously interpreted, avoided in nested comparisons, and were not tested in the current study (Markon, 2019).

Clarifying whether a single factor (and hence single total score) reflects the factor structure of the SCS is important for more accurate and valid measure scoring. Unclear also is which model is superior *relative* to other posited models. We compared nested models of six different factor structures of the SCS, avoiding bifactor models. We predicted a six-factor (Petrocchi et al., 2014) or a hierarchical model with six lower-order factors loading on to two higher-order factors (Montero-Marín et al., 2016) would fit best.

Method

Participants

We report how we determined our sample size, all data exclusions (if any), all data inclusion/exclusion criteria, whether inclusion/exclusion criteria were established prior to data analysis, all measures in the study, and all analyses. We report absolute and comparative model fit indices and exact p-values (see Data Analyses section). A power analysis showed a sample of N=818 is required to detect a small effect (.15; α=.05; 80% power), estimating the most complex model in the current study (Soper, 2021). With the only eligibility criteria being participants were in their first- or second-year, Canadian undergraduates (N=1158, 72.8% women, 26.6% men, .6% non-binary; *M[SD]* age=19.0[*2.3*], n=157 excluded for missing data) completed a survey in fall 2016 and winter 2017 as part of a larger study (Strickland et al., 2019).

Procedure

All first-and second-year students were emailed invitations to complete a 30-minute online survey (compensated with gift cards or course credits), with three weekly reminders (response rate: 30.4%).

Measures

Participants completed the 26-item SCS (Neff, 2003b), using a 5-point frequency scale (1=never to 5=almost always), including three positively oriented and three negatively oriented facets.

Data analyses

Lower-order one, two (self-caring, self-coldness), three (corresponding to the three positive/negative self-compassion components; Neff 2003a), and six-factor models were tested. Hierarchical models tested included one higher-order global self-compassion factor with six lower-order factors, and two higher-order factors (self-caring, self-coldness) with six lower-order factors.

Models were tested using the lavaan package in R using the diagonally weighted least squares (DWLS) estimator for ordinal items. Fit indices and standard errors used the robust variation. The following absolute fit indices suggest adequate fit: Standardized Root Mean Square Residual (SRMR)<.08, Comparative Fit Index (CFI) and Tucker Lewis Index (TLI)>.90, and Root Mean Square Error of Approximation (RMSEA)<.06 (Hu & Bentler, 1999). Nested models were compared with likelihood ratio tests. While the code and full output from all models are accessible (see Strickland, 2021), the raw data needed to reproduce all of the reported results are not due to ethical constraints.

Results

The lower-order one-factor model fit indices were poor: $X^2(299, N=1129)=12966$, $p<.001$, SRMR=.14, RMSEA=.19, CFI=.78, TLI=.76. Standardized factor loadings ranged from .43-.74. The lower-order two-factor model showed acceptable fit on absolute fit indices and poor fit on relative fit indices: $X^2(298, N=1129)=2245$, $p<.001$, SRMR=.06, RMSEA=.08, CFI=.90,

TLI=.89. Standardized factor loadings ranged from .54-.77 for self-caring and .62-.82 for self-coldness. The factors were significantly intercorrelated (r=-.46, p<.001; SE=0.02). The three-factor model fit poorly: X^2(296, N=1129)=12669, p<.001, SRMR=.14, RMSEA=.13, CFI=.79, TLI=.77. Standardized factor loadings ranged from .61-.77 for self-kindness vs. self-judgement, .45-.76 for mindfulness vs. overidentification, and .45-.75 for common humanity vs. isolation. Latent variables were strongly related, ranging from r=.86, p<.001; SE=0.02 (self-kindness, common humanity) to r=.93, p<.001; SE=0.02 (mindfulness, common humanity). The correlated six-factor structure produced a non-positive definite covariance matrix. Inspection of the eigenvalues suggested that the overidentification factor had a negative eigenvalue and large latent correlations with other factors (i.e.,>.90) suggesting it is non-unique/inseparable from the other factors.

The hierarchical model involving one higher-order factor and six lower-order factors showed poor fit: X (293, N=1129)=9303, p<.001, SRMR=.12, RMSEA=.12, CFI=.82, TLI=.80. At the lower-order level, standardized factor loadings ranged from .71-.79 for self-kindness, .55-.79 for mindfulness, .60-.78 for common humanity, .66-.83 for self-judgement, .63-.87 for overidentification, and .66-.83 for isolation. The lower-order factors showed standardized factor loadings from .64-.88 on the self-compassion factor. The hierarchical model with two higher-order factors (self-caring, self-coldness) and six lower-order factors (three facets/higher-order factor) showed adequate fit across indices: X^2(292, N=1129)=1728, p<.001, SRMR=.05, RMSEA=.07, CFI=.92, TLI=.91 (Figure 1).

The following models were compared to the hierarchical two-factor model using likelihood ratio tests: one-factor (X^2(7)=176.5, p<.001), two-factor (X^2(6)=239.7, p<.001), three-

factor ($X^2(4)=99.3$, $p<.001$), and one higher-order and six lower-order factors $X^2(7)=22.1$, $p<.001$). In all cases, the hierarchical two-factor model fit best.

Discussion

Several models of the SCS were compared using nested model comparisons. Contrary to Neff (2003b), a one-factor structure of the SCS was not supported. Few have replicated Neff's (2003b) findings and recent studies did not test the two-factor hierarchical model (Neff et al., 2019). A general factor may therefore be an inappropriate and invalid representation of the SCS (e.g., Brenner et al., 2017, Cleare et al., 2018; Williams et al., 2014), and use of a single SCS total score is not recommended. Rather, nested model comparisons suggest a plausible two-factor structure, with the more complex hierarchical two-factor model showing slight improvement over the lower order two-factor model, potentially because it better captured measure flaws (e.g., cross-or low-loading items). The two higher-order factors appear to represent self-coldness and self-caring, each explained by their corresponding three subfactors (e.g., self-coldness: self-judgement, isolation, and overidentification). The SCS should be analyzed with structural equation modelling using the higher- and lower-order factors rather than summed scores as predictors/outcomes.

The finding that the SCS is better reflected by two related factors than one general factor matches Gilbert's (2010) conceptualization of self-compassion as a two-process construct, supported by the moderate correlation between the higher-order self-caring and self-coldness factors. Moreover, the hierarchical two-factor structure advances findings supporting a two-factor lower-order structure: Costa et al. (2016) and Lopez et al. (2015) undertook modifications before finding good fit indices, suggesting higher-order domains with lower-order facets. Brenner's (2017) results also suggest the hierarchical two-factor model is a better alternative

than a one factor higher-or lower-order model, accounting for the difficulty interpreting nested bifactor model comparison (Markon, 2019).

Our difficulty estimating the six-factor model has been reported in previous studies (e.g., Lopez et al., 2015). Notably, Coroiu et al. (2018) found a six-factor structure showed good fit indices when items were randomly assigned to each factor, suggesting six factors may be an unreliable representation of the SCS. Moreover, Williams et al. (2014) supported a six-factor structure but did not compare the relative fit of a hierarchical two-factor structure.

The large sample and ordinal scoring are strengths, but our undergraduate sample may restrict generalizability of our study. The non-unique factor in the six-factor model and the high lower-order factor loadings in the hierarchical model (e.g.,>.95) suggests redundant items. Additionally, there is room for improvement on the fit of the two-factor hierarchical model, a result potentially driven by poorly specified items. The differences favoring the hierarchical model were also relatively slight and the lower-order two factor model is an alternative generally well-fitting structure. Future research might explore the factor structure of the 12-item short-form SCS (Raes et al., 2011). Altogether, nested comparisons informed the selection of the two-factor hierarchical model as the better-fitting relative factor structure of the SCS.

References

Bonifay, W., & Cai, L. (2017). On the complexity of item response theory models. *Multivariate Behavioral Research, 52*(4), 465-484. doi:10.1080/00273171.2017.1309262

Bonifay, W., Lane, S.P., & Reise, S.P. (2017). Three concerns with applying a bifactor model as a structure of psychopathology. *Clinical Psychological Science, 5*(1), 184-186. doi:10.1177/2167702616657069

Brenner, R.P., Vogel, D., Credé, M., Terence J.G., & Kivilghan, D.M. (2017). Two is more valid than one: Examining the factor structure of the Self-Compassion Scale. *Journal of Counseling Psychology, 64,* 696-707. doi:10.1037/cou0000211

Castilho, P., Pinto-Gouveia, J., & Duarte, J. (2015). Evaluating the multifactor structure of the long and short versions of the Self-Compassion Scale in a clinical sample. *Journal of Clinical Psychology, 71,* 856-870. doi:10.1002/jclp.22187

Cleare, S., Gumley, A., Cleare, C., & O'Connor, R. (2018). An investigation of the factor structure of the Self-Compassion Scale. *Mindfulness, 9,* 618-628. doi:10.1007/s12671-017-0803-1

Coroiu, A., Kwakkenbos, L., Moran, C., Thombs, B., Albani, C., Bourkas, S…. Körner, A. (2018). Structural validation of the Self-Compassion Scale with a German general population sample. *PLoS ONE, 13,* E0190771. doi:10.1371/journal.pone.0190771

Costa, J., Marôco, J., Pinto-Gouveia, J., Ferreira, C., & Castilho, P. (2016). Validation of the psychometric properties of the Self-Compassion Scale: Testing the factorial validity and factorial invariance of the measure among borderline personality disorder, anxiety disorder, eating disorder and general populations. *Clinical Psychology & Psychotherapy, 23,* 460-468. doi:10.1002/cpp.1974

DiStefano, C., Zhu, M., & Mindrila, D. (2009). Understanding and using factor scores: Considerations for the applied researcher. *Practical Assessment, Research, and Evaluation, 14*, 1-11. doi:10.7275/da8t-4g52

García-Campayo, J., Navarro-Gil, M., Andrés, E., Montero-Marin, J., López-Artal, L., & Demarzo, M. (2014). Validation of the Spanish versions of the long (26 items) and short (12 items) forms of the Self-Compassion Scale. *Health and Quality of Life Outcomes, 12*, 1-9. doi:10.1186/1477-7525-12-4

Gilbert, P. (2010). *Compassion focused therapy: Distinctive features*. Routledge.

Hu, L., & Bentler, P. (1999). Cut-off criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1-55. doi:10.1080/10705519909540118

Kotsou, I., & Leys, C. (2016). Self-Compassion Scale: Psychometric properties of the French translation and its relations with psychological well-being, affect and depression. *PLoS One, 11*, 1-11. doi:10.1371/journal.pone.0152880

Lopez, A., Sanderman, R., Smink, A., Zhang, Y., Van Sonderen, E., Ranchor, A., & Schroevers, M. (2015). A reconsideration of the Self-Compassion Scale's total score: Self-compassion versus self-criticism. *PloS One, 10,* 1-12. doi:10.1371/journal.pone.0132940

Markon, K.E. (2019). Bifactor and hierarchical models: Specification, inference, and interpretation. *Annual Review of Clinical Psychology, 15,* 51-69. doi:10.1146/annurev-clinpsy-050718-095522

Montero-Marín, J., Gaete, J., Demarzo, M., Rodero, B., Lopez, L.C.S., & García-Campayo, J. (2016). Self-criticism: A measure of uncompassionate behaviors toward the self, based

on the negative components of the Self-Compassion Scale. *Frontiers in Psychology, 7,* 1281. doi:10.3389/fpsyg.2016.01281

Murray, A.L., & Johnson, W. (2013). The limitations of model fit in comparing the bi-factor versus higher-order models of human cognitive ability structure. *Intelligence*, *41*, 407-422. doi:10.1016/j.intell.2013.06.004

Neff, K.D. (2003a). Self-compassion: An alternative conceptualization of a healthy attitude toward oneself. *Self and Identity, 2,* 85-101. doi:10.1080/15298860309032

Neff, K.D. (2003b). The development and validation of a scale to measure self-compassion. *Self and Identity, 2*, 223-250. doi:10.1080/15298860309027

Neff, K.D., Tóth-Király, I., Yarnell, L.M., Arimitsu, K., Castilho, P., Ghorbani, N.… Mantzios, M. (2019). Examining the factor structure of the Self-Compassion Scale in 20 diverse samples: Support for use of a total score and six subscale scores. *Psychological Assessment, 31,* 27-45. doi:10.1037/pas0000629

Neff, K.D., Whittaker, T., & Karl, A. (2017). Examining the factor structure of the Self-Compassion Scale in four distinct populations: Is the use of a total scale score justified? *Journal of Personality Assessment, 99,* 596-607. doi:10.1080/00223891.2016.1269334

Petrocchi, N., Ottaviani, C., & Couyoumdjian, A. (2014). Dimensionality of self-compassion: Translation and construct validation of the Self-Compassion Scale in an Italian sample. *Journal of Mental Health, 23,* 72-77. doi:10.3109/09638237.2013.841869

Raes, F., Pommier, E., Neff, K.D., & Van Gucht, D. (2011). Construction and factorial validation of a short form of the Self-Compassion Scale. *Clinical Psychology & Psychotherapy, 18,* 250-255. doi:10.1002/cpp.702

Reise, S.P., Kim, D.S., Mansolf, M., & Widaman, K.F. (2016). Is the bifactor model a better model or is it just better at modeling implausible responses? Application of iteratively reweighted least squares to the Rosenberg Self-Esteem Scale. *Multivariate Behavioral Research, 51,* 818-838. doi:10.1080/00273171.2016.1243461

Soper, D. S. (2021). A priori sample size calculator for structural equation models [Software]. Available from https://www.danielsoper.com/statcalc

Strickland, N.J. (2021). Clarifying the factor structure of the Self-Compassion Scale: Nested comparisons of six confirmatory factor analysis models [Data Set]. https://doi.org/10.5683/SP2/BEIDFZ

Strickland, N.J., Wekerle, C., Kehayes, I.-L., Thompson, K., Dobson, K., & Stewart, S.H. (2019). Self-compassion as a compensatory resilience factor for the negative emotional outcomes of alcohol-involved sexual assault among undergraduates. *International Journal of Child and Adolescent Resilience, 6,* 52–69. doi:10.7202/1069076ar

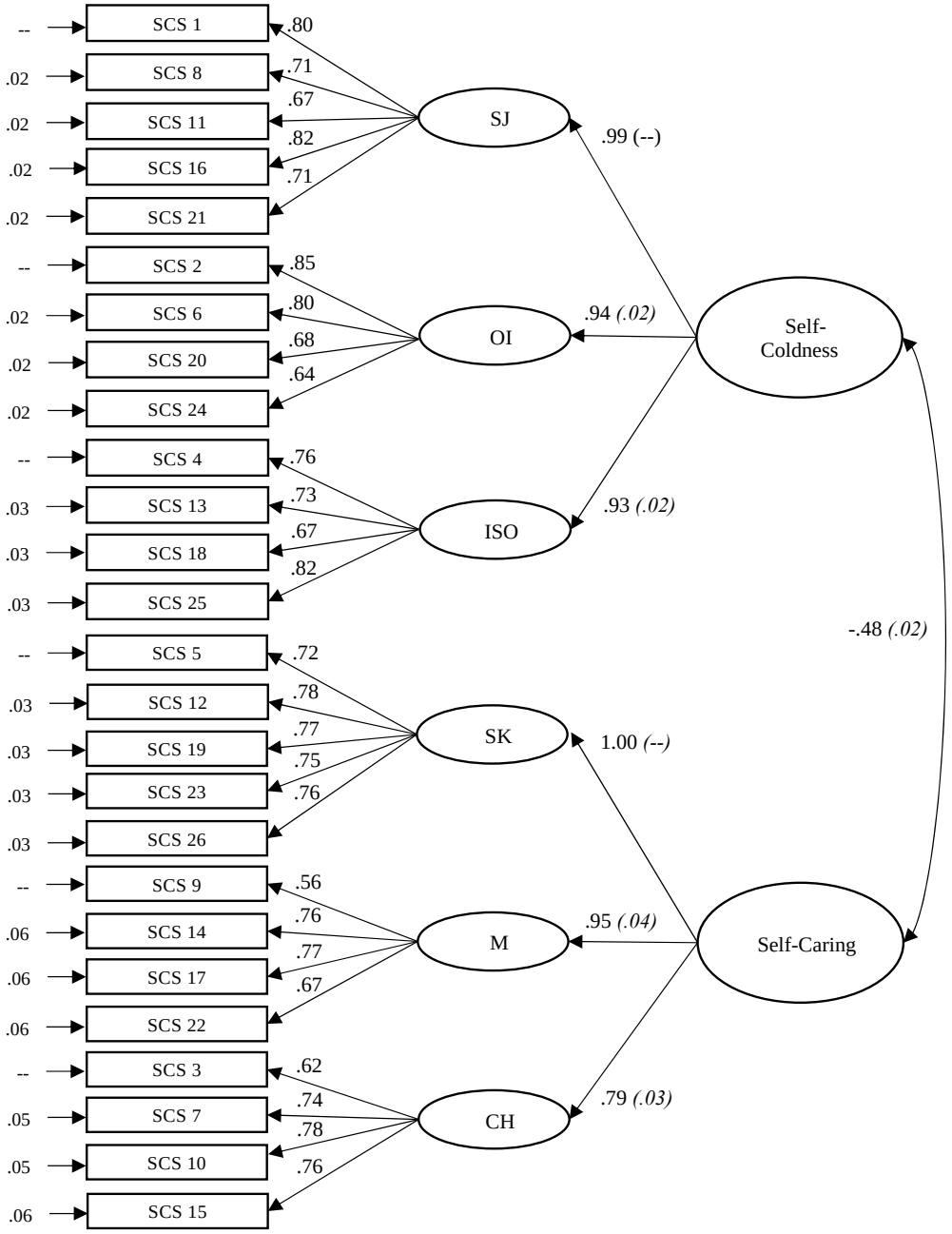
Tóth-Király, I., Bőthe, B., & Orosz, G. (2017). Exploratory structural equation modeling analysis of the Self-Compassion Scale. *Mindfulness, 8,* 881-892. doi:10.1037/met0000045

Williams, M., Dalgleish, T., Karl, A., Kuyken, W., & Reynolds, C.R. (2014). Examining the factor structures of the Five Facet Mindfulness Questionnaire and the Self-Compassion Scale. *Psychological Assessment, 26,* 407-418. doi:10.1037/a0035566

Zeller, M., Yuval, K., Nitzan-assayag, Y., & Bernstein, A. (2015). Self-compassion in recovery following potentially traumatic stress. *Journal of Abnormal Child Psychology, 43,* 645-653. doi:10.1007/s10802-014-9937-y

Zeng, X., Wei, J., Oei, T.P., & Liu, X. (2016). The Self-Compassion Scale is not validated in a Buddhist sample. *Journal of Religion and Health, 55,* 1996-2009. doi:10.1007/s10943-016-0205-z

Figure 1. Two-factor hierarchical model of the Self-Compassion Scale with standardized factor loadings (*p<.001). Standardized coefficients are reported with standard errors *(SE)* and error terms are left of items. SJ=self-judgement, OI=overidentification, ISO=isolation, SK=self-kindness, M=mindfulness, and CH=common humanity. All factor loadings and correlations were significant (p<.001).



Open Science

Open Data: Raw data needed to reproduce all of the reported results are not openly accessible due to ethical constraints, however, the code and output are accessible (Strickland, 2021).

Open Materials: I confirm that there is sufficient information for an independent researcher to reproduce all of the reported methodology, including the code and results output (Strickland, 2021).

Preregistration of Studies and Analysis Plans: This study was not preregistered with an analysis plan.