







Shortening the Relationship Disillusionment Scale to a Three-Item Measure (RDS-3)

Carson R. Dover¹ , Alan Reifman¹ , Mozhgan Rezvani Shakib¹ ,

Aubrey S. M. Pickett¹ , C. Rebecca Oldham² , Sylvia Niehuis¹ 

[1] *Human Development & Family Sciences, Texas Tech University, Lubbock, TX, USA.* [2] *Department of Human Sciences, Middle Tennessee State University, Murfreesboro, TN, USA.*

Measurement Instruments for the Social Sciences, 2026, Vol. 8, Article e20841, <https://doi.org/10.5964/miss.20841>

Received: 2025-11-12 • **Accepted:** 2025-12-15 • **Published (VoR):** 2026-02-13

Handling Editor: Matthias Ziegler, Humboldt Universität zu Berlin, Berlin, Germany

Corresponding Author: Sylvia Niehuis, 1301 Akron Ave, Lubbock, TX 79415, USA. E-mail: sylvia.niehuis@ttu.edu

Supplementary Materials: Code [see [Index of Supplementary Materials](#)]



Abstract

Relationship disillusionment—perceiving that one’s romantic relationship is not meeting expectations and getting worse and viewing one’s spouse/partner in an increasingly negative light—has been shown to predict likelihood of relationship breakup, independent of relationship satisfaction, commitment, and duration. Current versions of the Relationship Disillusionment Scale (RDS), the leading measure in this area, have 11 or more items. The present research examined properties (factor loadings, internal consistency and test-retest reliability, convergent and concurrent criterion validity, and dyadic correlations) of full-length RDS measures in three datasets to identify a small number of items that could comprise a psychometrically sound short form of the RDS. A three-item RDS version emerged, which performed comparably well to full-length versions. Hence, we recommend the RDS-3 for use in future research when survey length is an issue (e.g., daily diary studies; surveys taken on mobile phones).

Keywords

measurement, relationship disillusionment, romantic relationships, scale development, shortened measures

Romantic relationships play a crucial role in many individuals’ lives, promoting happiness and meaning. Relationship problems, however, can shatter partners’ worldviews and lead to breakup or divorce (Niehuis, Reifman, & Oldham, 2019; Niehuis et al., 2021). A key challenge for romantic partners is the mismatch between expectations (which



are often unrealistic or lofty) and reality, a phenomenon central to the Disillusionment Model (e.g., [Huston et al., 2001](#); [Niehuis et al., 2011](#)). Capturing its multifaceted nature, [Niehuis, Reifman, and Oldham \(2019\)](#) define disillusionment as “a perception or feeling that experiences with one's partner and relationship have not gone nearly as well as one had expected and that restoring the relationship to a satisfactory state may be hopeless” (p. 467).

Niehuis and colleagues ([Niehuis, 2007](#); [Niehuis & Bartell, 2006](#)) developed the 16-item Marital Disillusionment Scale (MDS), which demonstrated high internal consistency and convergent and discriminant validity. In revising the scale (now called the Relationship Disillusionment Scale [RDS] to encompass dating and cohabiting relationships), several items were removed and others added, yielding an 11-item instrument ([Niehuis et al., 2015](#); [Niehuis et al., 2021](#)). Such changes stemmed partly from some items involving pointed judgments (e.g., “I no longer really like my partner as a person,” “My relationship is no longer as important to me as it used to be”), with which respondents may have been reluctant to agree due to social desirability. The RDS has successfully predicted self-reported breakup likelihood among married and cohabiting couples (controlling for relationship satisfaction, commitment, and duration) in a national sample ([Niehuis et al., 2015](#)).

Like any instrument, the RDS can be improved further. Longer surveys can yield lower participation, greater participant burden, skipped questions, or even abandonment of the survey ([Burgard et al., 2020](#); [Couper & Mavletova, 2015](#); [Eisele et al., 2022](#); [Gibson & Bowling, 2020](#); [Trull & Ebner-Priemer, 2020](#)). The increasing use of mobile devices to take surveys adds other complications, as small screens and expectations for quicker interactions make long questionnaires less practical ([Link et al., 2014](#)). Therefore, short measures may work better when brevity is essential (e.g., diary studies).

Scale shortening depends on how conceptually multifaceted a construct is. Some have argued constructs that are singularly focused, unidimensional, homogeneous, etc., may be sufficiently measured with as few as one item ([Allen et al., 2022](#)). Self-esteem is one such construct ([Robins et al., 2001](#)). Other constructs, however, may incorporate multiple, somewhat interrelated, facets. In our view, disillusionment falls into the latter category, involving perceptions of unmet expectations for one's partner and relationship, perceived relationship deterioration, and increasingly unfavorable views of one's spouse/partner. Although the RDS often exhibits alpha/internal consistency of $> .90$ (e.g., [Niehuis et al., 2015](#)), high alpha does not necessarily indicate unidimensionality (e.g., [Morera & Stokes, 2016](#)). Accordingly, we propose that a shortened disillusionment scale would need at least three items.

Relationship scholars stand to benefit from using brief, effective measurements. Historically, important constructs in relationship sciences have included love, satisfaction, conflict, and commitment ([Joel et al., 2020](#); [Rubin, 1970](#)). [Niehuis et al. \(2024\)](#) evaluated the psychometric properties of brief measures of these four constructs compared to

full-length versions. They found full-length scales for love, satisfaction, conflict, and commitment could indeed be shortened to single-item measures and still show acceptable psychometric qualities (see also [Hendrick et al.'s, 1998](#), shortening of the Love Attitudes Scale).

Psychometric theory is important in framing efforts to shorten measures. Scholars have argued that more items may promote reliability, but may also lead to item-specific errors that, if non-random, compound each other ([Robins et al., 2001](#)). Other research by [Thalmayer and colleagues \(2011\)](#) further explained how shorter measures are inherently a compromise between convenience and reliability/predictive validity. This compromise between scale length and reliability is important because scholars have noted how reliability (specifically test-retest reliability) is crucial for testing validity ([McCrae et al., 2011](#)). Hence, longer scales carry both risk and reward and scales of varying numbers of items (i.e., full-length and shortened versions) should be thoroughly tested before employing them in research.

Methods: Instrument and Scale Development

To select items for a shortened RDS, we first conducted Confirmatory Factor Analyses (CFA) using every possible combination of three items from the original 11-item scale. After the most psychometrically promising sets of three items were observed, we evaluated long and short versions on reliability (internal consistency, test-retest) and validity (convergent between shortened and longer RDS versions; criterion-related between the RDS-3 and relationship satisfaction). We utilized three datasets (one national and two with local university students), enhancing generalizability of the research. Our studies will contribute an efficient tool for assessing disillusionment to advance research and clinical practice in this area ([Prouty et al., 2024](#)).

Data and Participants

We used three datasets that included a full-length (11-item) RDS version. The datasets and their designs (cross-sectional or longitudinal; dyadic or individual) are described below.

National Sample

The first dataset was the Married and Cohabiting Couples 2010 project ([National Center for Family and Marriage Research \[NCFMR\], 2010](#)). Investigators sought a nationally representative U.S. sample of heterosexual married and cohabiting couples ages 18-64 (i.e., randomly sampling from a larger address-based panel). The final sample, however, underrepresented Black and Hispanic respondents and overrepresented highly educated ones. The sample contained 752 married and 323 cohabiting couples.

Undergraduate Sample 1 (Predictors of Disillusionment [POD] Study)

The next sample came from an online (Qualtrics) study of predictors (e.g., recalled family-of-origin experiences) of disillusionment in young adults. Participants ($N = 674$ individuals who had been in at least one romantic relationship, with 6% married) were recruited from a southwestern U.S. university via an extra-credit participation pool. This sample was not divided by gender due to a heavy imbalance (75% women).

Undergraduate Sample 2 (College Dating Study)

The final sample was from an online (Qualtrics) three-wave study of unmarried, heterosexual dating couples (assessments one month apart) at the same southwestern university. Participants were recruited via the university's e-mail announcement system. Sample sizes at the three waves were 121, 84, and 72 couples, respectively.

Overall Demographics

Men and women were evenly matched in the dyadic studies, but as noted, not in the POD study. White respondents ranged from 80% in the national sample to around 65% in the two university samples. Hispanic representation ranged from 21% in both of the college samples to 8% in the national sample. Black representation ranged from 2–8%, with other race/ethnicity categories comprising 5–8% of the samples. College attendees/graduates comprised about 71% of the national study.

Measures

Disillusionment

The 11-item RDS (scored 1-5 from *Strongly Disagree* to *Strongly Agree*; Niehuis et al., 2015) was used in each of the samples. High scores represented greater disillusionment (requiring score reversals in the NCFMR survey). Our short-form development began by using all 11 items that appeared in the three main datasets.

Relationship Satisfaction

Relationship satisfaction was assessed via multiple measures, with items being recoded as necessary so that higher scores indicated greater satisfaction. In the NCFMR study, we used two different scales for greater generalizability. The first scale was a latent variable (within the CFAs) consisting of 4 items: "Taking all things together, how satisfied are you with your relationship with your spouse or partner?", "How satisfied are you with how well your spouse/partner listens to you?" (both measured from 1 = *Very satisfied* to 5 = *Very dissatisfied*), "My spouse/partner shows love and affection toward me," and "My spouse/partner encourages me to do things that are important to me" (both measured from 1 = *Strongly agree* to 5 = *Strongly disagree*). We created a z-score for each item before including them in the final models. The second scale was a one-item measure that

asked: “How would you rate your relationship with your current spouse/partner?” and was measured from 1 (*Completely unhappy*) to 10 (*Completely happy*).

The POD study assessed satisfaction via three items from the Relationship Assessment Scale (RAS; Hendrick, 1988): “How well does/did your partner/spouse meet your needs?”, “In general, how satisfied are/were you with your relationship?”, and “How good is/was your relationship compared to most?” (1 = *Not at All* to 5 = *Extremely*; $\alpha = .90$). Convergent validity of the RAS was demonstrated by its correlation with other satisfaction measures (Renshaw et al., 2011).

Finally, we used the MOQ satisfaction scale (Huston et al., 1986) in the online college dating study. Participants were asked: “How do you experience your relationship?” They were given a list of bipolar word-pairs (e.g., miserable-enjoyable, interesting-boring) on an anchored scale from 1 to 7. The items were reverse-coded as needed so that higher scores reflected more positive experiences in the relationship. The 11 items were averaged together to create a mean score (women’s $\omega = .95$, men’s $\omega = .93$).

Analysis Plan

For the initial item-selection step, we used the NCFMR study because it had the largest and most representative sample. CFA—using every possible three-item combination appearing in the datasets—served as our initial gatekeeping mechanism. Within the CFA approach, we optimized model fit, reliability, and convergent validity. Previous research has utilized a similar approach when shortening established scales (Steger & Schütz, 2025). Because the data consisted of matched couples, we ran the analyses twice—once for women and once for men. The mathematical formula for *n choose k*, or *11 choose 3* in the present context, yields 165 possible combinations of three items, with the three items being loaded onto a latent variable in the respective models. Versions of the three-item latent RDS were then correlated with both the one- and four-item satisfaction measures within the model. After running the models, fit measures (Comparative Fit Index [CFI] and Root Mean Square Error of Approximation [RMSEA]) were extracted, as well as McDonald’s omega (ω , for reliability), factor loadings for each of the three items in each model, and correlations between the variables (for validity). After the CFA identified the most promising items for a shortened disillusionment scale, we analyzed measurement invariance between men and women. This has been highlighted as an important step in dyadic romantic relationship research (Sakaluk et al., 2021). Finally, we examined the RDS-3 for reliability (internal consistency and test-retest), and validity (convergent and concurrent criterion-related) using the other datasets. We used a variety of R packages to run the analyses throughout the study (Hester & Bryan, 2024; Jorgensen et al., 2022; Makowski et al., 2022; R Core Team, 2024; Revelle, 2024; Rosseel, 2012; Schauburger & Walker, 2025; Wickham, Averick, et al., 2019; Wickham & Henry, 2025; Wickham, Miller, & Smith, 2023)¹.

Quality Criteria

Confirmatory Factor Analyses

The initial set of 165 models in the NCFMR data were first run for women. The first set of criteria was good model fit ($CFI > .949$, $RMSEA < .08$), leaving 127 possible combinations. Next, we examined only models that had one variable for each of the three disillusionment facets, leaving 36 models. From these 36 models, the one with the highest reliability ($\omega = .928$) was with the following items: “Our relationship has changed for the worse,” “I am very disappointed in my marriage/relationship,” and “I feel no longer quite as positively about my spouse/partner as I once did.” All factor loadings were $> .894$. The RDS-3 also correlated significantly and strongly with the four-item relationship satisfaction scale ($-.88$, $p < .001$) and the one-item satisfaction measure ($-.84$, $p < .001$) in the expected negative direction. The process was similar for men. Filtering for good model fit left 152 combinations, and models with one variable from each facet yielded 39 possible models. The model with the highest reliability ($\omega = .923$) contained the same items as for women. The factor loadings were all $> .895$. The measure also correlated strongly and significantly with the four-item ($-.88$, $p < .001$) and one-item ($-.81$, $p < .001$) satisfaction measures. Based on these consistent results showing both good reliability and validity, we determined that the three aforementioned items would comprise the RDS-3.

Affirming that the RDS-3, although correlating substantially with relationship satisfaction, retained the ability to independently predict key outcomes, we showed in the NCFMR data that the RDS-3 significantly predicted men and women’s self-rated breakup likelihood, controlling for relationship satisfaction, commitment, and duration. This finding replicates a result with the full RDS in the [Niehuis et al. \(2015\)](#) study.

Dyadic Validity

Measurement Invariance

We again utilized the NCFMR dataset to establish measurement invariance between men and women. Model fit was unable to be established for the configural model (i.e., group analysis) because the three-item CFA was just-identified and had 0 degrees of freedom. When constraining the factor loadings to be equal between groups for the metric model, the model did not fit significantly worse than the configural model ($\Delta\chi^2[2] = 0.219$, $p = .896$; $\Delta CFI = .000$), establishing metric measurement invariance between men and women. Finally, when constraining the intercepts and factor loadings to be the same across groups, the scalar model did fit significantly worse ($\Delta\chi^2[2] = 6.995$, $p = .03$; $\Delta CFI =$

1) The script for analyses and a link to the national data set can be found in the Supplementary Materials (see [Dover, 2025](#)).

-.001), demonstrating that there is not scalar measurement invariance between men and women.

Dyadic Correlations

Using the dyadic datasets (NCFMR and college dating) we analyzed how the RDS-3 (compared to the “long version” comprised of the other eight items not used) correlated between romantic partners. In the NCFMR study, married husbands and wives correlated at $r = .54$ [95% Confidence Interval: .49, .59] on the RDS-3 and at $r = .57$ [.52, .61] on the long version. The correlations for cohabiting couples were similarly strong, with $r = .60$ [.53, .67] on the RDS-3 and $r = .56$ [.48, .63] on the long version. All these NCFMR correlations were significant ($p < .001$). At Phase 1 in the college dating survey, the RDS-3 correlated between partners at $r = .51$ [.36, .63] and the longer scale at $r = .54$ [.40, .66]. Overall, these results demonstrate that the RDS-3 has similar dyadic correlations to that of the long (eight-item) scale.

Convergent Validity

To examine convergent validity, we correlated the RDS-3 with the long RDS (i.e., the eight remaining items) in the same datasets. For married men in the NCFMR study, they correlated at $r = .90$ [.89, .92] and for married women they correlated at $r = .92$ [.91, .93]. For cohabiting men, the scales correlated at $r = .93$ [.91, .94], and at $r = .93$ [.91, .94] for cohabiting women. At Phase 1 of the college dating study, the RDS-3 correlated with the long scale at $r = .86$ [.81, .90] for men and $r = .88$ [.83, .92] for women. In the predictors of disillusionment study, the RDS-3 and long (eight-item) scale correlated $r = .94$ [.93, .95]. All of the correlations for convergent validity meet Allen et al.'s (2022) criterion for “good” or “excellent” convergent validity.

Concurrent Criterion Validity

To test concurrent criterion validity in each dataset, we correlated available relationship satisfaction measures with both the RDS-3 and long (8-item) RDS. In the POD study, the RDS-3 correlated with the three-item relationship assessment scale at $r = -.51$ [-.56, -.44], while the long version correlated with this same scale at $r = -.52$ [-.58, -.46]. For men at Phase 1 of the college dating study, the RDS-3 correlated with the MOQ scale $r = -.77$ [-.83, -.68], and the long scale correlated at $r = -.85$ [-.90, -.79]. For women at Phase 1, the RDS-3 correlated with the MOQ scale $r = -.83$ [-.88, -.76], and the full-length scale correlated $r = -.86$ [-.90, -.80]. The key finding was that relationship satisfaction showed comparably strong correlations (in a negative direction) with the short and long RDS versions.

Internal Consistency Reliability

To test internal-consistency reliability, we calculated McDonald's omega (ω). In the NCFMR study, married women and married men both had excellent reliability ($\omega = .93$ and $.92$, respectively). Cohabiting women and men had similarly excellent reliability ($\omega = .93$ and $.92$, respectively). In the POD study, the sample also had excellent reliability ($\omega = .94$). In the college dating sample, both men ($\omega = .83$) and women ($\omega = .91$) had good reliability. The RDS-3 demonstrated very good internal consistency across all of our samples.

Test-Retest Reliability

We examined test-retest reliability of the RDS-3 using the college dating sample (with a one-month interval between phases). For men, the RDS-3 was correlated strongly between Phases 1 and 2 ($r = .58$; [.41, .71]), Phases 2 and 3 ($r = .72$; [.58, .82]), and Phases 1 and 3 ($r = .65$; [.47, .77]). A similar pattern emerged for women between Phases 1 and 2 ($r = .68$; [.54, .78]), 2 and 3 ($r = .76$; [.64-.84]), and 1 and 3 ($r = .57$; [.39-.71]).

Discussion

We examined the psychometric properties of Marital/Relationship Disillusionment Scale items in multiple datasets to help develop a short version for future studies (the RDS-3). As we noted, relationship disillusionment is a multifaceted construct, incorporating expectations, fulfillment (or not) thereof, a temporal sense of things getting worse, and hopelessness at repairing the relationship. Hence, we expected multiple items (although fewer than in full-length versions) to be necessary to represent the construct well. Based on our CFA results, we selected three items—pertaining to the belief that one's romantic relationship is not meeting expectations, and that both the relationship and one's view of their spouse/partner is getting worse—that capture much of this conceptualization. These items correspond well to areas of the brain that activated in a previous study while participants were in an fMRI scanner, in response to reminders of disillusioning (vs. dissatisfying) relationship events in participants' lives (Niehuis, Reifman, Al-Khalil, et al., 2019). The activated brain areas to disillusionment in the fMRI experiment corresponded to evaluation, reflection, and reconciling conflicting information, which in the present context pertain to references in the RDS-3 items about conflicting information (i.e., high expectations vs. changes for the worse; initially favorable vs. currently negative impressions of one's marriage/relationship and spouse/partner). Further supporting our shortened scale's ability to reflect the larger concept, we found extremely high correlations (around .90) between the three-item version and the remaining disillusionment items. The newly developed short form also exhibited the expected strong negative

correlations with measures of relationship satisfaction. Finally, the RDS-3 also displayed high internal-consistency and good test-retest reliability.

Our study featured several strengths, as well as limitations. On the positive side, we drew from three different datasets that, in the aggregate, included married, cohabiting, and dating couples, and represented populations ranging from community to national (within the United States). Further, our tests for concurrent criterion-related validity used four different measures of relationship satisfaction, bolstering the generalizability of those findings. Finally, we used a longitudinal dataset, allowing examination of test-retest reliability. A limitation of the scale was that the test-retest reliability was not as strong as expected. The time between phases was rather short, so the correlations were expected to be larger. Scholars have noted the importance of test-retest reliability for scales (McCrae et al., 2011), especially how it relates to the validity of the scale. Research should prioritize more longitudinal work on the nature of the RDS-3 and how it changes. Despite this limitation, the scale demonstrated strong reliability and validity across many domains, demonstrating its ability to be used in future work.

The Marital/Relationship Disillusionment Scale has been empirically successful in predicting breakup likelihood with rigorous controls for possible confounders (Niehuis et al., 2015), which we replicated using the RDS-3 instead of the 11-item version used in the previous study. Research in the future could focus also on testing the scale's validity in predicting actual breakups. Because disillusionment rises longitudinally in response to partners' relationship transgressions (Niehuis, Reifman, & Oldham, 2019), it is a good "barometer" of relationship functioning. Further, the study of adults' current relationship disillusionment in the context of family-of-origin factors offers extensive therapeutic implications (Prouty et al., 2024). We thus conclude that the RDS-3 is a viable option for scholars to include if space in their questionnaire is limited, as well as for assessment in clinical contexts.

Funding: The national *Married and Cohabiting Couples Study* is supported by the National Center for Family and Marriage Research, which is funded by a cooperative agreement, grant number 5 U01 AE000001-04, between the Assistant Secretary for Planning and Evaluation (ASPE) in the United States Department of Health and Human Services (HHS) and Bowling Green State University.

Acknowledgments: We thank Randi Black and Annie Boyles for providing feedback on an earlier draft of this article and all the undergraduate and graduate students who have assisted on our lab projects over the years.

Competing Interests: The authors have declared that no competing interests exist.

Ethics Statement: Community samples were collected under IRB approval from Texas Tech University. The NCFMR dataset was obtained from the public domain.

Data Availability: A link to the national data set is included in the Supplementary Materials (see [Dover, 2025](#)).

Supplementary Materials

The Supplementary Materials include the script for analyses and a link to the national data set (see [Dover, 2025](#)).

Index of Supplementary Materials

Dover, C. (2025). *Shortening the Relationship Disillusionment Scale to a Three-Item Measure (RDS-3)* [Script]. OSF. <https://osf.io/zfnh7>

References

- Allen, M. S., Iliescu, D., & Greiff, S. (2022). Single item measures in psychological science. *European Journal of Psychological Assessment, 38*(1), 1–5. <https://doi.org/10.1027/1015-5759/a000699>
- Burgard, T., Bošnjak, M., & Wedderhoff, N. (2020). Response rates in online surveys with affective disorder participants: A meta-analysis of study design and time effects between 2008 and 2019. *Zeitschrift für Psychologie, 228*(1), 14–24. <https://doi.org/10.1027/2151-2604/a000394>
- Couper, M. P., & Mavletova, A. (2015). A meta-analysis of breakoff rates in mobile web surveys. In D. Toninelli, R. Pinter, & P. de Pedraza (Eds.), *Mobile research methods: Opportunities and challenges of mobile research methodologies* (pp. 81–98). Ubiquity Press. <https://doi.org/10.5334/bar.f>
- Eisele, G., Vachon, H., Lafit, G., Kuppens, P., Houben, M., Myin-Germeys, I., & Viechtbauer, W. (2022). The effects of sampling frequency and questionnaire length on perceived burden, compliance, and careless responding in experience sampling data in a student population. *Assessment, 29*(2), 136–151. <https://doi.org/10.1177/1073191120957102>
- Gibson, A. M., & Bowling, N. A. (2020). The effects of questionnaire length and behavioral consequences on careless responding. *European Journal of Psychological Assessment, 36*(2), 410–420. <https://doi.org/10.1027/1015-5759/a000526>
- Hendrick, C., Hendrick, S. S., & Dicke, A. (1998). The Love Attitudes Scale: Short form. *Journal of Social and Personal Relationships, 15*(2), 147–159. <https://doi.org/10.1177/0265407598152001>
- Hendrick, S. S. (1988). A generic measure of relationship satisfaction. *Journal of Marriage and the Family, 50*(1), 93–98. <https://doi.org/10.2307/352430>
- Hester, J., & Bryan, J. (2024). glue: Interpreted string literals. R package version 1.8.0, <https://CRAN.R-project.org/package=glue>
- Huston, T. L., Caughlin, J. P., Houts, R. M., Smith, S. E., & George, L. J. (2001). The connubial crucible: Newlywed years as predictors of marital delight, distress, and divorce. *Journal of Personality and Social Psychology, 80*(2), 237–252. <https://doi.org/10.1037/0022-3514.80.2.237>
- Huston, T. L., McHale, S., & Crouter, A. (1986). “When the honeymoon’s over”: Changes in the marriage relationship over the first year. In S. Duck & R. Gilmour (Eds.), *The emerging field of personal relationships* (pp. 109–132). Erlbaum.
- Joel, S., Eastwick, P. W., Allison, C. J., Arriaga, X. B., Baker, Z. G., Bar-Kalifa, E., Bergeron, S., Birnbaum, G. E., Brock, R. L., Brumbaugh, C. C., Carmichael, C. L., Chen, S., Clarke, J., Cobb, R.

- J., Coolsen, M. K., Davis, J., de Jong, D. C., Debrot, A., DeHaas, E. C., . . . Wolf, S. (2020). Machine learning uncovers the most robust self-report predictors of relationship quality across 43 longitudinal couples studies. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(32), 19061–19071. <https://doi.org/10.1073/pnas.1917036117>
- Jorgensen, T. D., Pornprasertmanit, S., Schoemann, A. M., & Rosseel, Y. (2022). semTools: Useful tools for structural equation modeling. R package version 0.5-6. Retrieved from <https://CRAN.R-project.org/package=semTools>
- Link, M. W., Murphy, J., Schober, M. F., Buskirk, T. D., Hunter Childs, J., & Langer Tesfaye, C. (2014). Mobile technologies for conducting, augmenting and potentially replacing surveys: Executive summary of the AAPOR task force on emerging technologies in public opinion research. *Public Opinion Quarterly*, *78*(4), 779–787. <https://doi.org/10.1093/poq/nfu054>
- Makowski, D., Wiernik, B. M., Patil, I., Lüdtke, D., & Ben-Shachar, M. S. (2022). correlation: Methods for correlation analysis (0.8.3) [R package]. <https://CRAN.R-project.org/package=correlation> (Original work published 2020)
- McCrae, R. R., Kurtz, J. E., Yamagata, S., & Terracciano, A. (2011). Internal consistency, retest reliability, and their implications for personality scale validity. *Personality and Social Psychology Review*, *15*(1), 28–50. <https://doi.org/10.1177/1088868310366253>
- Morera, O. F., & Stokes, S. M. (2016). Coefficient α as a measure of test score reliability: Review of 3 popular misconceptions. *American Journal of Public Health*, *106*(3), 458–461. <https://doi.org/10.2105/AJPH.2015.302993>
- National Center for Family and Marriage Research. (2010). Married and cohabiting couples. *Inter-university Consortium for Political and Social Research*. <https://doi.org/10.3886/ICPSR31322.v1>
- Niehuis, S. (2007). Convergent and discriminant validity of the Marital Disillusionment Scale. *Psychological Reports*, *100*(1), 203–207. <https://doi.org/10.2466/pr0.100.1.203-207>
- Niehuis, S., Adamczyk, K., Trepanowski, R., Celejewska, A., Ganclerz, M., Frydrysiak, A., Reifman, A., & Willis-Grossmann, E. (2021). Development of the Polish-Language Relationship Disillusionment Scale and its validation. *Personal Relationships*, *28*(4), 1017–1041. <https://doi.org/10.1111/pere.12394>
- Niehuis, S., & Bartell, D. (2006). The Marital Disillusionment Scale: Development and psychometric properties. *North American Journal of Psychology*, *8*(1), 69–83.
- Niehuis, S., Davis, K., Reifman, A., Callaway, K., Luempert, A., Oldham, C. R., Head, J., & Willis-Grossmann, E. (2024). Psychometric evaluation of single-item relationship satisfaction, love, conflict, and commitment measures. *Personality & Social Psychology Bulletin*, *50*(3), 387–405. <https://doi.org/10.1177/01461672221133693>
- Niehuis, S., Lee, K.-H., Reifman, A., Swenson, A., & Hunsaker, S. (2011). Idealization and disillusionment in intimate relationships: A review of theory, method, and research. *Journal of Family Theory & Review*, *3*(4), 273–302. <https://doi.org/10.1111/j.1756-2589.2011.00100.x>
- Niehuis, S., Reifman, A., Al-Khalil, K., Oldham, C. R., Fang, D., O’Boyle, M., & Davis, T. H. (2019). Functional magnetic resonance imaging activation in response to prompts of romantically disillusioning events. *Personal Relationships*, *26*(2), 209–231. <https://doi.org/10.1111/pere.12272>

- Niehuis, S., Reifman, A., & Lee, K. H. (2015). Disillusionment in cohabiting and married couples: A national study. *Journal of Family Issues*, 36(7), 951–973. <https://doi.org/10.1177/0192513X13498594>
- Niehuis, S., Reifman, A., & Oldham, C. R. (2019). Effects of relational transgressions on idealization of and disillusionment with one's partner: A three-wave longitudinal study. *Personal Relationships*, 26(3), 466–489. <https://doi.org/10.1111/perc.12287>
- Prouty, A. M., Niehuis, S., Reifman, A., & Willis-Grossmann, E. (2024). Young adults' romantic disillusionment as a function of family-of-origin dynamics. *Journal of Couple & Relationship Therapy*, 23(1), 40–61. <https://doi.org/10.1080/15332691.2023.2267168>
- R Core Team. (2024). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Renshaw, K. D., McKnight, P., Caska, C. M., & Blais, R. K. (2011). The utility of the Relationship Assessment Scale in multiple types of relationships. *Journal of Social and Personal Relationships*, 28(4), 435–447. <https://doi.org/10.1177/0265407510377850>
- Revelle, W. (2024). *psych: Procedures for psychological, psychometric, and personality research* [R package] (Version 2.4.6). <https://CRAN.R-project.org/package=psych>
- Robins, R. W., Hendin, H. M., & Trzesniewski, K. H. (2001). Measuring global self-esteem: Construct validation of a single-item measure and the Rosenberg Self-Esteem Scale. *Personality & Social Psychology Bulletin*, 27(2), 151–161. <https://doi.org/10.1177/0146167201272002>
- Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48(2), 1–36. <https://doi.org/10.18637/jss.v048.i02>
- Rubin, Z. (1970). Measurement of romantic love. *Journal of Personality and Social Psychology*, 16(2), 265–273. <https://doi.org/10.1037/h0029841>
- Schauberger, P., & Walker, A. (2025). *openxlsx: Read, write and edit xlsx files* [R package] (Version 4.2.8). <https://CRAN.R-project.org/package=openxlsx>
- Sakaluk, J. K., Fisher, A. N., & Kilshaw, R. E. (2021). Dyadic measurement invariance and its importance for replicability in romantic relationship science. *Personal Relationships*, 28(1), 190–226. <https://doi.org/10.1111/perc.12341>
- Steger, D., & Schütz, A. (2025). Assessing trust in science: Development and validation of a short scale for adolescents and adults. *Measurement Instruments for the Social Sciences*, 7, Article e16225. <https://doi.org/10.5964/miss.16225>
- Thalmayer, A. G., Saucier, G., & Eigenhuis, A. (2011). Comparative validity of brief to medium-length Big Five and Big Six Personality Questionnaires. *Psychological Assessment*, 23(4), 995–1009. <https://doi.org/10.1037/a0024165>
- Trull, T. J., & Ebner-Priemer, U. W. (2020). Ambulatory assessment in psychopathology research: A review of recommended reporting guidelines and current practices. *Journal of Abnormal Psychology*, 129(1), 56–63. <https://doi.org/10.1037/abn0000473>
- Wickham H., Averick M., Bryan J., Chang W., McGowan L.D., François R., Grolemund G., Hayes A., Henry L., Hester J., Kuhn M., Pedersen T.L., Miller E., Bache S.M., Müller K., Ooms J., Robinson D., Seidel D.P., Spinu V., Takahashi K., Vaughan D., Wilke C., Woo K., & Yutani H. (2019).

Welcome to the tidyverse. *Journal of Open Source Software*, 4(43), Article 1686.

<https://doi.org/10.21105/joss.01686>

Wickham, H., & Henry, L. (2025). *purrr: Functional programming tools* [R package] (Version 1.1.0).

<https://CRAN.R-project.org/package=purrr>

Wickham, H., Miller, E., & Smith, D. (2023). *haven: Import and export 'SPSS', 'Stata' and 'SAS' files* [R

package] (Version 2.5.4). <https://CRAN.R-project.org/package=haven>