INTERNATIONALE HOCHSCHULE

www.iu.de

IU DISCUSSION PAPERS

Sozialwissenschaften

PSYCHOMETRIC CEILING EFFECTS IN THE ZIP

WHEN LOVE DOES NOT DIFFERENTIATE

CONSTANTIN SCHUBART
DAVID MEIßNER

IU Internationale Hochschule

Juri-Gagarin-Ring 152

99084 Erfurt

Telefon: +49 421.166985.23

Fax: +49 2224.9605.115

Kontakt/Contact: kerstin.janson@iu.org

Autorenkontakt/Contact to the author(s):

Prof. Dr. Constantin Schubart
constantin.schubart@iu.org
ORCID-ID: 0009-0008-9259-0533 (Open Researcher and Contributor ID)

David Meißner
david.meissner.research@gmx.net
ORCID-ID: 0009-0009-9516-6292 (Open Researcher and Contributor ID)

IU Internationale Hochschule
Juri-Gagarin-Ring 152
99084 Erfurt
Email: constantin.schubart@iu.org

IU Discussion Papers, Reihe: Sozialwissenschaften, Vol. 6, No. 12 (NOV 2026)

ISSN: 2750-0675

DOI: https://doi.org/10.56250/4158

Webseite: https://repository.iu.org

PSYCHOMETRIC CEILING EFFECTS IN THE ZIP

When Love Does Not Differentiate

Constantin Schubart
David Meißner

ABSTRACT:

The German ZIP (Zufriedenheit in Paarbeziehungen) is an established seven-item measure of relationship satisfaction, but little is known about the item-level performance of its direct love item in highly satisfied samples. This exploratory post hoc study examined whether the ZIP love item shows a ceiling effect and whether a reduced six-item score excluding this item retains comparable psychometric properties. Data came from 685 German-speaking individuals in heterosexual romantic relationships lasting at least six months. Item descriptives, corrected item-rest correlations, internal consistency estimates, confirmatory factor analyses, and manifest score comparisons were used to compare the original seven-item ZIP (ZIP-7) with a reduced six-item variant (ZIP-6). The love item showed a severe ceiling effect: 94.31% of participants selected the maximum response category, with the highest mean (M = 4.92), lowest variability (SD = 0.35), strongest negative skewness (−4.80), and highest kurtosis (23.99) among all ZIP items. In the continuous-item ML CFA for ZIP-7, the love item had the lowest standardized factor loading (λ = .507), lowest explained variance (R^2 = .257), and highest standardized residual variance (θ = .743). The separate ML CFA for ZIP-6 yielded CFI = .987, TLI = .978, RMSEA = .067, and SRMR = .021. Cronbach's alpha differed only slightly between ZIP-7 (α = .881) and ZIP-6 (α = .887). ZIP-7 and ZIP-6 were almost perfectly correlated, r = .997, and showed virtually identical associations with intimacy satisfaction, both rounded to r = .46. The findings support examining ZIP-6 as an exploratory sensitivity score requiring further validation, particularly in highly satisfied samples.

KEYWORDS:

relationship satisfaction; ZIP; love item; ceiling effects; psychometrics

AUTOR:INNEN



Dr. Constantin Schubart *ist seit 2020 Professor für Allgemeine Betriebswirtschaftslehre an der IU Internationale Hochschule im Dualen Studium am Standort Erfurt. Seine Schwerpunkte liegen im Bereich Managerial Economics. Er verfügt über mehr als 20 Jahre Berufserfahrung in Lehre, Forschung und Praxis.*



David Meißner *hat einen Bachelorabschluss in Psychologie und studiert derzeit im Master Psychologie. Zudem arbeitet er als wissenschaftlicher Mitarbeiter. Seine wissenschaftliche Tätigkeit umfasst ein breites Spektrum psychologischer Themen mit interdisziplinären Bezügen.*

Introduction

RELATIONSHIP SATISFACTION AND ITS MEASUREMENT

Relationship satisfaction is a central construct in relationship research, linked to relational outcomes such as relationship stability and divorce-related outcomes (Bühler & Orth, 2025; Plopa et al., 2023), as well as individual outcomes such as physical and mental health (Kiecolt-Glaser & Newton, 2001; Robles, 2014). Given this relevance, valid, reliable, and economical measures are needed. The Relationship Assessment Scale (RAS) is an internationally established brief measure of global relationship satisfaction (Hendrick, 1988), and its German adaptation, the ZIP (*Zufriedenheit in Paarbeziehungen*), is commonly used in German-speaking research (Hassebrauck, 1991). The ZIP comprises seven items covering partner need fulfillment, global relationship satisfaction, comparison with other couples, relationship-related regret, expectation fulfillment, love for the partner, and relationship problems. It has been described as a reliable measure with a clear one-factor structure (Hassebrauck, 1991). Item-level differentiation in the present sample, however, also depends on how respondents use the available response categories. More generally, rating scales are among the most frequently used instruments in social science data collection, and their design has been studied in relation to psychometric quality criteria such as reliability and validity. In particular, the number of response categories determines the degree of differentiation of a rating scale along the intended continuum (Menold & Bogner, 2016). Accordingly, item-level response distributions are not merely descriptive but provide relevant information about whether an item differentiates respondents within the relevant range of the construct.

RESPONSE BIAS AND COGNITIVE CONSISTENCY MECHANISMS

Direct love items may be problematic from a measurement perspective because they ask respondents to report on a socially valued and normatively expected aspect of romantic relationships, which may make maximum endorsement especially likely and thereby limit the item's use of the available response categories. Self-reports can be affected by socially desirable responding, particularly for socially valued attitudes or behaviors (Crowne & Marlowe, 1960). This does not imply that relationship satisfaction reports are mere social-desirability artifacts; prior work suggests that the influence of general socially desirable responding on such reports is small (Visschers et al., 2017). The relative contributions of these explanations cannot be inferred from a ceiling effect alone.

The present concern is narrower and item-specific: a direct love item may be especially prone to maximum endorsement because it asks respondents to affirm a normatively expected emotional commitment that is closely tied to their self-understanding as a committed partner. Moreover, the love item differs from the remaining ZIP items in its attributional emphasis. Although all ZIP items involve subjective evaluation, the ZIP item content suggests that low endorsement of the love item may be harder to externalize because it directly concerns one's own emotional attachment to the partner, whereas low endorsement of the other items can more readily be attributed to partner behavior, relationship functioning, unmet expectations, comparisons with other relationships, or relationship problems (Hassebrauck, 1991). The proposed attributional differences are hypotheses derived from the item content and were not tested in the present study.

From a cognitive dissonance perspective, reporting limited love while remaining in a relationship may create inconsistency, especially if respondents believe that one should be with a partner one truly loves. Such inconsistency may be reduced by maintaining a coherent and positively idealized relationship evaluation, consistent with accounts of dissonance reduction (Festinger, 1957; Lavergne & Pelletier, 2016). Separately, research on positive illusions suggests that satisfied partners often idealize their relationship (Murray & Holmes, 1997). Thus, maximum endorsement of a love item may reflect genuine satisfaction, normative self-presentation, cognitive consistency, and motivated idealization simultaneously, limiting its ability to differentiate among highly satisfied respondents.

CROSS-LINGUISTIC EQUIVALENCE OF THE LOVE ITEM

Although the ZIP love item corresponds to the RAS item assessing how much respondents love their partner, *love* and *Liebe* may differ in their pragmatic implications. Cross-cultural research shows that "I love you" varies substantially in its contextual range across cultures (Wilkins & Gareis, 2006). In the United States–Germany comparison, verbal love expression appears less central in Germany, where "Ich liebe dich" has traditionally been more narrowly associated with private, meaningful, and committed romantic contexts (Gareis & Wilkins, 2011). These cross-cultural findings motivate a hypothesis about language-related response processes, but do not establish that German wording produces a ceiling effect in the ZIP love item.

PSYCHOMETRIC IMPLICATIONS: CEILING EFFECTS AND RESTRICTED ITEM VARIANCE

These mechanisms become psychometrically relevant when responses accumulate at the upper end of a scale and limit differentiation among respondents with high construct levels. A commonly used criterion defines ceiling effects as present when more than 15% of respondents select the highest possible score (Terwee et al., 2007). This criterion was proposed in the context of health status questionnaires and is used here as a descriptive reference rather than as a universal threshold for judging the psychometric adequacy of individual items. Such effects imply restricted observed variance; range restriction can attenuate or distort associations with other variables (Sackett & Yang, 2000). Thus, an item may be theoretically meaningful but provide little discriminative information when most respondents endorse the maximum category, because item-level validity also depends on differentiation at relevant levels of the latent construct (Yang & Kao, 2014). Moreover, ceiling and floor effects can bias estimates, increase uncertainty, and affect statistical inference (Šimkovic & Träuble, 2019).

Method

The study followed the Declaration of Helsinki. Written informed consent was obtained electronically from all participants. The present study did not include experimental manipulations. Anonymized data, the codebook, JASP analysis outputs, and the Electronic Supplementary Material are available on OSF (Meißner, 2026).

Data were collected through an online questionnaire implemented in SoSci Survey. Participants were included if they were currently in a heterosexual romantic relationship and provided valid ZIP responses. Exclusions were based on failed attention checks, incomplete responses, unusually fast completion, inconsistent relationship information, or not meeting the relationship-status criteria for the present analyses.

The analytic sample comprised 685 German-speaking participants in heterosexual romantic relationships lasting at least six months, including 478 women (69.78%) and 207 men (30.22%). Participants ranged in age from 18 to 80 years and were predominantly younger adults; 78.39% were between 18 and 30 years old, 16.79% were between 31 and 50 years old, and 4.82% were older than 50 years. Relationship duration ranged from 6 months to more than 30 years: 7.59% reported relationships lasting 6 to 12 months, 32.85% 1 to 3 years, 27.15% 3 to 5 years, 21.17% 5 to 10 years, 5.99% 10 to 20 years, 3.50% 20 to 30 years, and 1.75% more than 30 years.

MEASURE

Relationship satisfaction was assessed with the German ZIP (Hassebrauck, 1991). The love item was presented in German as “Wie sehr lieben Sie Ihren Partner?”. The verbal endpoint labels were “sehr” and “gar nicht”. Responses were scored so that higher values indicated greater relationship satisfaction. The ZIP was administered using a five-point response format, consistent with Dinkel and Balck (2005), rather than the original seven-point format used by Hassebrauck (1991). The number of response categories may influence response distributions and the extent of observed ceiling effects. Analyses focused on the love item and compared the original seven-item score (ZIP-7) with a reduced six-item score excluding this item (ZIP-6).

ANALYTIC STRATEGY

Analyses were conducted as an exploratory post hoc item-level evaluation of the ZIP. The focus on the love item emerged during initial descriptive inspection, where responses showed a pronounced concentration at the maximum scale value; therefore, the analyses should be interpreted as exploratory rather than confirmatory. Item descriptives were computed for all ZIP items, including means, standard deviations, skewness, kurtosis, and the percentage of maximum responses. Descriptive subgroup analyses examined whether the love-item ceiling effect was present across gender, marital status, and shared religious practice. Finally, the original seven-item ZIP and a reduced six-item ZIP were compared using corrected item-rest correlations, Cronbach’s alpha, McDonald’s omega, and one-factor confirmatory factor analyses. The main CFA models were estimated using maximum likelihood estimation. Model fit was evaluated using χ^2, CFI, TLI, RMSEA, and SRMR. As a robustness check, separate CFA models were estimated using DWLS for ordered categorical indicators.

Results

DESCRIPTIVE ITEM CHARACTERISTICS

Descriptive analyses indicated generally high relationship satisfaction across ZIP items, with all items showing negative skewness and no missing values (Table 1).

Table 1

Item Descriptive Statistics

Item	*M*	*SD*	Skewness	Kurtosis	% maximum response	λ ZIP-7	R^2	Residual variance
fulfillment	4.18	0.74	-0.97	1.46	33.72	.758	.575	.425
satisfaction	4.48	0.70	-1.47	2.44	57.08	.841	.707	.293
comparison	4.46	0.75	-1.34	1.34	58.83	.745	.555	.445
regret_r	4.52	0.70	-1.49	2.26	61.46	.760	.578	.422
expectation	4.24	0.84	-1.10	1.00	44.67	.799	.639	.361
love	4.92	0.35	-4.80	23.99	94.31	.507	.257	.743
problems_r	3.73	0.78	-1.02	1.10	9.49	.632	.400	.600

Note. ZIP = Zufriedenheit in Paarbeziehungen. *M* = mean; *SD* = standard deviation. % maximum response refers to the percentage selecting the highest response category. λ ZIP-7 = standardized loading in the seven-item model. Reverse-scored items are indicated by _r.

The love item showed the clearest distributional restriction: 646 of 685 participants endorsed the maximum response category, corresponding to 94.31% (*M* = 4.92, *SD* = 0.35). It also had the highest mean, lowest variability, strongest negative skewness, and highest kurtosis among all ZIP items.

Exploratory subgroup analyses showed that this ceiling effect was highly consistent across gender, marital status, and shared religious practice, with maximum-response rates ranging from 91.91% to 97.33%. Full subgroup descriptives for the love item are shown in the Electronic Supplementary Material.

MODEL COMPARISON

At the manifest score level, the reduced six-item score showed a slightly lower mean but greater variability than the original seven-item score (ZIP-7: *M* = 4.36, *SD* = 0.54, Mdn = 4.57; ZIP-6: *M* = 4.27, *SD* = 0.60, Mdn = 4.50). The reduced score also showed a somewhat wider observed range (ZIP-7: 2.00–5.00; ZIP-6: 1.67–5.00) and slightly less pronounced skewness and kurtosis (ZIP-7: skewness = −1.55, kurtosis = 2.74; ZIP-6: skewness = −1.44, kurtosis = 2.41). The manifest ZIP-7 and ZIP-6 average scores were almost perfectly correlated, Pearson's $r = .997$, $p < .001$, and Spearman's $\rho = .999$, $p < .001$. Participants' relative ordering was therefore very similar in this sample. However, the scores shared six items, so their high correlation does not establish equivalent construct validity. Separate one-factor CFA models were estimated for the seven-item and six-item scores. The ML model for ZIP-7 yielded $\chi^2(14) = 52.31$, $p < .001$, CFI = .983, TLI = .975, RMSEA = .063, 90% CI [.046, .082], and SRMR = .025. In this continuous-item ML model, the love item had the lowest standardized loading, λ = .507, 95% CI [.447, .567], and explained variance, $R^2 = .257$, and the highest standardized residual variance, θ = .743. The ML model for ZIP-6 yielded $\chi^2(9) = 36.39$, $p < .001$, CFI = .987, TLI = .978, RMSEA = .067, 90% CI [.045, .090], and SRMR = .021. Both models showed acceptable approximate fit when evaluated separately. Their fit indices do not establish that deleting the love item improves measurement or preserves construct validity. Table 2

presents the model-specific and score-level summaries. Complete standardized loadings, residual variances, and their 95% confidence intervals for both ML models are available in OSF output 02_CFA_ML_standardized_estimates.html.

Table 2

Score and Model Comparison of ZIP-7 and ZIP-6

Model	Item	*M*	*SD*	Range	CFI	TLI	RMSEA	SRMR	ω	α	Lowest item-rest r
ZIP-7	7 items	4.36	0.54	2.00–5.00	.983	.975	.063	.025	.893	.881	.473
ZIP-6	6 items	4.27	0.60	1.67–5.00	.987	.978	.067	.021	.889	.887	.604

Note. CFA statistics refer to the continuous-item ML models. Reliability estimates refer to the frequentist reliability analysis. Fit indices describe each model separately and are not tests of equivalence between the two scores.

The love item had the lowest corrected item-rest correlation in ZIP-7, $r = .473$; correlations for the other ZIP-7 items ranged from .600 to .779. Cronbach's alpha was .881, 95% CI [.866, .895], for ZIP-7 and .887, 95% CI [.871, .903], for ZIP-6. McDonald's omega was .893, 95% CI [.880, .905], and .889, 95% CI [.876, .902], respectively. Thus, alpha increased slightly whereas omega decreased slightly after item deletion. These descriptive differences do not establish a meaningful improvement in reliability. Correlations with the single-item intimacy-satisfaction measure were $r = .4625$ for ZIP-7 and $r = .4621$ for ZIP-6, both $p < .001$. These very similar point estimates describe one relationship-related criterion in this sample; they do not establish equivalent criterion or construct validity.

The supplementary ordinal DWLS analysis yielded $\chi^2(14) = 22.05$, $p = .078$, CFI = .9993, TLI = .9990, RMSEA = .029, 90% CI [.000, .051], and SRMR = .027 for ZIP-7, and $\chi^2(9) = 14.32$, $p = .111$, CFI = .9995, TLI = .9992, RMSEA = .029, 90% CI [.000, .057], and SRMR = .021 for ZIP-6. The love item's standardized DWLS loading was .837, 95% CI [.760, .913], with $R^2 = .700$ and standardized residual variance $\theta = .300$. Thus, its low loading in the continuous-item ML analysis was not reproduced in the ordinal analysis. The concentration of observed responses at the maximum category remained, but the factor-loading evidence depended on the estimation approach.

Discussion

PSYCHOMETRIC IMPLICATIONS

The results showed a severe ceiling effect for the love item: 94.31% of participants selected the highest response category, far exceeding the commonly used 15% criterion for ceiling effects (Terwee et al., 2007). This distributional restriction was accompanied by the lowest variability, strongest negative skewness, and highest kurtosis among all ZIP items. In the continuous-item ML model, the love item also had the weakest factor loading, lowest explained variance, and highest standardized residual variance. Moreover, the pattern was consistent across subgroups defined by gender, marital status, and shared religious practice. Thus, in the present highly satisfied sample, the love item provided limited differentiation, particularly at the upper end of relationship satisfaction. Omitting the love item produced small descriptive differences in selected reliability estimates, score ordering, and the association

with intimacy satisfaction. The CFA models were evaluated separately, and their fit indices do not establish equivalent measurement properties. The pronounced ceiling effect may reflect genuinely high and relatively homogeneous levels of love in this sample. The present data cannot distinguish genuine homogeneity from the effects of item wording, response format, or other response processes. Removing the love item may also narrow the content coverage of the scale, even when selected reliability estimates and score associations change little. These findings should not be interpreted as evidence that the six-item version is generally superior to the established ZIP. ZIP-6 may be examined as an exploratory sensitivity score, while its content coverage and construct validity require further investigation.

INTEGRATION WITH EXISTING LITERATURE

Previous validation research supports distinguishing overall scale quality from the observed variability of individual items. Dinkel and Balck (2005), using Hassebrauck's German item wording with five response categories, found one dominant principal component, high internal consistency (α = .89), and lower RAS scores among participants in distressed relationships. Their love item was also strongly endorsed. Our ZIP-7 reliability estimate (α = .881) is consistent with this general pattern, while the present analysis focuses on restricted observed variability of a specific item.

Funk and Rogge (2007) demonstrated that measurement information for most instruments assessing relationship satisfaction decreased sharply at high satisfaction levels. Adamczyk et al. (2022) similarly found that the Polish RAS measured low to average satisfaction more precisely than high satisfaction. The present ceiling effect is compatible with these findings, but does not itself constitute an IRT-based assessment of measurement precision.

Fülöp et al. (2022) selected and validated the global satisfaction item as RAS-1. Although several external associations were similar to those of the full RAS, they reported a stronger association with love for the full scale. Their findings highlight the importance of considering item content and independent validation criteria when evaluating shortened scores. Their validation of RAS-1 does not establish that removing the love item preserves the content coverage or construct validity of ZIP-6.

The incremental contribution of the present study is therefore the documentation of an exceptionally pronounced ceiling effect for the love item (94.31% maximum responses) in a German-speaking sample of 685 participants, together with the descriptive consequences of omitting that item. The findings document limited observed differentiation in this sample; they do not establish general item inadequacy or validate a replacement scale.

THEORETICAL INTERPRETATION

The findings are consistent with the theoretical rationale outlined in the Introduction. A direct love item may be especially vulnerable to maximum endorsement because it asks respondents to affirm a normatively expected emotional commitment that may be harder to externalize than broader evaluations of relationship functioning. Reporting less than maximum love while remaining in a relationship may create perceived inconsistency, consistent with cognitive dissonance accounts (Festinger, 1957; Lavergne & Pelletier, 2016). Thus, maximum endorsement may reflect genuine affection and relationship satisfaction, but may also be reinforced by normative self-presentation, cognitive consistency, and motivated positive relationship perception (Crowne & Marlowe, 1960; Murray & Holmes, 1997).

The German wording may further contribute to this pattern, as *Liebe* and *Ich liebe dich* are typically associated with strong, committed romantic contexts (Gareis & Wilkins, 2011; Wilkins & Gareis, 2006). In a highly satisfied German-speaking sample, the ZIP love item may therefore function less as a fine-grained indicator of relationship satisfaction and more as an affirmation of basic emotional commitment. However, this interpretation remains tentative: the present study did not directly measure social desirability, cognitive dissonance, relationship idealization, or language-specific response processes, nor did it directly compare the German ZIP with the English RAS. Therefore, the findings should be interpreted as consistent with, but not direct evidence for, the proposed social-cognitive and linguistic mechanisms.

LIMITATIONS

Several limitations should be considered. First, the analyses were exploratory and post hoc, as the focus on the love item emerged during initial descriptive inspection. Therefore, the findings require replication in independent samples.

Second, the sample was not representative and consisted predominantly of younger adults and women in German-speaking heterosexual romantic relationships lasting at least six months. It was also characterized by high relationship satisfaction, although high satisfaction was not an inclusion criterion. The severe ceiling effect may therefore be less pronounced in more distressed, newly formed, non-heterosexual, or otherwise more heterogeneous samples. Additionally, the use of a five-point rather than the original seven-point response format may have contributed to the observed ceiling effect by reducing the available response differentiation. However, the extent of this influence cannot be determined from the present data, as response formats were not experimentally compared.

Third, the proposed mechanisms, including social desirability, attributional processes, cognitive consistency, cognitive dissonance, positive relationship illusions, and language-specific pragmatic effects, were not measured directly. Consequently, the present study can identify an item-level psychometric pattern but cannot determine which psychological or linguistic processes produced it.

Fourth, the observed ceiling effect does not establish inadequate measurement, because it may also reflect genuinely high and relatively homogeneous levels of love in this sample. Deleting the love item may narrow the content represented by the scale, even when selected statistical indicators change little. Moreover, the love item's loading differed substantially between the continuous-item ML and ordinal DWLS models, limiting a general interpretation of the item as a weak indicator of relationship satisfaction.

Fifth, the CFA models used different indicator sets, precluding the ordinary nested-model difference test and a direct AIC/BIC comparison. Finally, the score comparisons involved substantial item overlap and only one external single-item criterion. They do not establish equivalent content coverage or performance in distressed samples, longitudinal change, or prediction of relationship outcomes.

Conclusions

The present exploratory item-level analysis showed that the ZIP love item was affected by a severe ceiling effect and contributed comparatively little discriminative information. This response concentration limited observed differentiation in the present sample. Omitting the item produced only small descriptive changes in selected reliability estimates, score ordering, and the association with intimacy satisfaction. These findings support inspecting item-level distributions and transparently reporting exploratory sensitivity analyses. These findings do not imply that the love item is invalid or that ZIP-7 should generally be replaced. Rather, they suggest that researchers using the ZIP should inspect item-level distributions and consider six-item sensitivity analyses when the love item shows pronounced ceiling effects, particularly in highly satisfied samples.

Electronic Supplementary Material

The Electronic Supplementary Material (Electronic_Supplementary_Material.zip) contains (a) the anonymized dataset and codebook, (b) JASP analysis outputs (01–04), (c) a supplementary table containing subgroup descriptive statistics, and (d) a README file providing a detailed description of all supplemental components. These materials are available via OSF: https://doi.org/10.17605/OSF.IO/PKMDC. Additional information necessary for clarification can be provided by the author upon reasonable request.

References

Adamczyk, K., Kleka, P., & Frydrychowicz, M. (2022). Psychometric functioning, measurement invariance, and external associations of the Relationship Assessment Scale in a sample of Polish adults. *Scientific Reports, 12*, 22157. https://doi.org/10.1038/s41598-022-26653-6

Bühler, J. L., & Orth, U. (2025). Terminal decline of satisfaction in romantic relationships: Evidence from four longitudinal studies. *Journal of Personality and Social Psychology*. Advance online publication. https://doi.org/10.1037/pspp0000551

Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology, 24*(4), 349–354. https://doi.org/10.1037/h0047358

Dinkel, A., & Balck, F. (2005). An evaluation of the German Relationship Assessment Scale. *Swiss Journal of Psychology, 64*(4), 259–263. https://doi.org/10.1024/1421-0185.64.4.259

Festinger, L. (1957). *A theory of cognitive dissonance.* Stanford University Press. https://doi.org/10.1515/9781503620766

Fülöp, F., Bőthe, B., Gál, É., Cachia, J. Y. A., Demetrovics, Z., & Orosz, G. (2022). A two-study validation of a single-item measure of relationship satisfaction: RAS-1. *Current Psychology, 41*(4), 2109–2121. https://doi.org/10.1007/s12144-020-00727-y

Funk, J. L., & Rogge, R. D. (2007). Testing the ruler with item response theory: Increasing precision of measurement for relationship satisfaction with the Couples Satisfaction Index. *Journal of Family Psychology, 21*(4), 572–583. https://doi.org/10.1037/0893-3200.21.4.572

Gareis, E., & Wilkins, R. (2011). Love expression in the United States and Germany. *International Journal of Intercultural Relations*, *35*(3), 307–319. https://doi.org/10.1016/j.ijintrel.2010.06.006

Hassebrauck, M. (1991). ZIP – Ein Instrumentarium zur Erfassung der Zufriedenheit in Paarbeziehungen. *Zeitschrift für Sozialpsychologie, 22*(4), 256–259.

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

Kiecolt-Glaser, J. K., & Newton, T. L. (2001). Marriage and health: His and hers. *Psychological Bulletin*, *127*(4), 472–503. https://doi.org/10.1037/0033-2909.127.4.472

Lavergne, K. J., & Pelletier, L. G. (2016). Why are attitude–behaviour inconsistencies uncomfortable? Using motivational theories to explore individual differences in dissonance arousal and motivation to compensate. *Motivation and Emotion, 40*(6), 842–861. https://doi.org/10.1007/s11031-016-9577-3

Meißner, D. (2026, June 29). *Psychometric ceiling effects in the ZIP: When love does not differentiate* [Data set, codebook, analysis outputs, and supplementary material]. OSF. https://doi.org/10.17605/OSF.IO/PKMDC

Menold, N., & Bogner, K. (2016). *Design of rating scales in questionnaires*. GESIS Survey Guidelines. GESIS – Leibniz Institute for the Social Sciences. https://doi.org/10.15465/gesis-sg_en_015

Murray, S. L., & Holmes, J. G. (1997). A leap of faith? Positive illusions in romantic relationships. *Personality and Social Psychology Bulletin, 23*(6), 586–604. https://doi.org/10.1177/0146167297236003

Plopa, M., Plopa, W., & Skuzińska, A. (2023). Grounds for divorce: The dynamics of marriage satisfaction as a predictor of divorce. *Advances in Cognitive Psychology, 19*(3), 283–296. https://doi.org/10.5709/acp-0399-7

Robles, T. F. (2014). Marital quality and health: Implications for marriage in the 21st century. *Current Directions in Psychological Science*, *23*(6), 427–432. https://doi.org/10.1177/0963721414549043

Sackett, P. R., & Yang, H. (2000). Correction for range restriction: An expanded typology. *Journal of Applied Psychology, 85*(1), 112–118. https://doi.org/10.1037/0021-9010.85.1.112

Šimkovic, M., & Träuble, B. (2019). Robustness of statistical methods when measure is affected by ceiling and/or floor effect. *PLoS ONE*, *14*(8), e0220889. https://doi.org/10.1371/journal.pone.0220889

Terwee, C. B., Bot, S. D. M., de Boer, M. R., van der Windt, D. A. W. M., Knol, D. L., Dekker, J., Bouter, L. M., & de Vet, H. C. W. (2007). Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of Clinical Epidemiology*, *60*(1), 34–42. https://doi.org/10.1016/j.jclinepi.2006.03.012

Visschers, J., Jaspaert, E., & Vervaeke, G. (2017). Social desirability in intimate partner violence and relationship satisfaction reports: An exploratory analysis. *Journal of Interpersonal Violence*, *32*(9), 1401–1420. https://doi.org/10.1177/0886260515588922

Wilkins, R., & Gareis, E. (2006). Emotion expression and the locution "I love you": A cross-cultural study. *International Journal of Intercultural Relations*, *30*(1), 51–75. https://doi.org/10.1016/j.ijintrel.2005.07.003

Yang, F. M., & Kao, S. T. (2014). Item response theory for measurement validity. *Shanghai Archives of Psychiatry*, *26*(3), 171–177. https://doi.org/10.3969/j.issn.1002-0829.2014.03.010