The Relationship Satisfaction scale – Psychometric properties

Espen Røysamb[1,2], Joar Vittersø[3] and Kristian Tambs[2]

1) University of Oslo, Department of Psychology, Oslo, Norway
2) Norwegian Institute of Public Health, Oslo, Norway
3) University of Tromsø, Department of Psychology, Tromsø, Norway

Correspondence: Espen Røysamb, Department of Psychology, University of Oslo, P.O. Box 1094 Blindern, NO-0317 Oslo, Norway
E-mail: espen.roysamb@psykologi.uio.no   Telephone: +47 94 29 51 77   Telefax: +47 22 84 50 01

ABSTRACT

The aim of this study was to establish the psychometric properties of the new Relationship Satisfaction (RS) scale. Two population based samples were used: The Norwegian Mother and Child Cohort Study (MoBa, N=117,178) and The Quality of Life study (N=347). Convergent and discriminant validity was investigated in relation to the Quality of Marriage Index (QMI), the Satisfaction With Life Scale (SWLS), Relationship Satisfaction of partner, Big Five personality traits (IPIP50) and future relationship dissolution. The full scale with ten items (RS10) and a short version with five items (RS5) showed good psychometric properties. The scale has high internal and test-retest reliability and high structural, convergent, and discriminant validity. Measurement invariance across gender was established. Additionally, predictive validity was evidenced by prediction of future relationship dissolution. We conclude that the RS scale is highly useful as a generic measure of global relationship satisfaction.

Most people live much of their lives in close relationships to others. Relations with spouses or partners represent a core aspect of social life. Such relationships might be sources of support, love, health and well-being, but may also represent a troubling factor in life (1,2). Divorce rates of 40-50%, as found in many countries (3-6), testify to the complexity of marriage and romantic relationships. In order to study the causes, correlates and consequences of relationship quality there is a need for reliable and valid measures of various aspects of relationships. The present study reports on the construction and psychometric properties of a brief new scale measuring generalized relationship satisfaction.

Several questionnaire scales have previously been developed in this field, including the Quality Marriage Index (7), the Dyadic Adjustment Scale (8), the Marital Satisfaction Inventory (9), the Kansas Marital Satisfaction Scale (10), the Marital Satisfaction and Commitment Scales (11), the Commitment Inventory (12), the Relationship Assessment Scale (13), the Marital Satisfaction Scale (14), the Couples Satisfaction Index (15) and the Interpersonal Reactivity Index for Couples (16). Such scales represent important contributions to measurement in this field. In addition to findings supporting the psychometric properties of the scales, several studies have shown the value of measuring relationship qualities in terms of predicting important outcomes (2,17).

Despite the existence of several established scales, for some purposes these scales have inherent limitations that imply a need for further development of measurements. First, several of the scales were constructed for use in therapy settings and are rather lengthy and thereby inadequate for use in large surveys.

Second, the construction and validation of the scales is partly based on limited samples or clinical groups. Third, some scales tap specific relational aspects, such as commitment, communication and support. Specific dimensions of relationship quality are important to address. Nevertheless, global measures of satisfaction are required for some research purposes. Fourth, given that gender roles have changed in recent decades and vary across cultures there is a need for measures with minimal culture- or time-specific content. Finally, most existing scales refer to marriage as the relationship to address. There is a need for scales that use broader terms, including both married couples, cohabiting couples and partners.

The current study was conducted within the framework of the Norwegian Mother and Child Cohort study (MoBa). This is a longitudinal population-based study recruiting women, and their partners, during pregnancy (18,19). The participants are followed up with continuously new waves of data collections. The questionnaires comprise substantial sections on psychological and social issues, and include measures such as life satisfaction, anxiety, depression, self-esteem, self-efficacy and life events, both on mothers, fathers and children within the same families. Within the context of the MoBa study a new measure of relationship satisfaction was developed.

The aim of this study was to develop and investigate the psychometric properties of the ten-item relationship satisfaction (RS10) scale and the five-item short version (RS5). More specifically, we aimed to examine the reliability and structural, concurrent and predictive validity of the full and short RS scale, and to suggest a cut-off point for dissatisfactory relationships.

METHODS

Materials/samples

Two separate samples were used. *Sample 1* is based on the Quality of Life study conducted in 2004. From a gross sample of 1500 Norwegians above 18 years old drawn randomly from the telephone directory, we were able to locate an adjusted gross sample of 1267 persons. A total of 466 persons participated, yielding a response rate of 36.78%. Out of these, 347 reported being in a relationship and responded to the relationship scales (see below). Age ranged from 19 to 71 years with a mean of 48 (SD = 12.5), and 57.4% of the sample were women. This sample was used for initial testing and validation of the RS scale.

Sample 2 consists of 117,178 participants (women and men) in the MoBa, conducted by the Norwegian Institute of Public Health (18), and approved by the Regional Committee for Medical Research and the Norwegian Data Inspectorate. Informed consent was obtained from each MoBa participant upon recruitment. MoBa is a prospective population-based pregnancy cohort study that started in 1999 and stopped recruiting new participants in 2008. The response rate was 40.6% for the whole MoBa study, 43.5% for the current sample. All spouses/partners of recruited women were also invited to complete one questionnaire, and 77% of the invited men agreed to participate. Pregnant women were recruited through a postal invitation in connection with a routine ultrasound examination offered all pregnant women at 17-18 weeks of gestation. After obtaining informed consent, the women completed the initial questionnaire (Time 1). Further questionnaires are completed later in pregnancy and in the months and years following birth. Details are described elsewhere (18). Also, it has been shown that although prevalence estimates are biased due to self-selection, estimates of exposure-outcome associations are generally not much affected (19).

The current study is based on version IV of the quality-assured data files released for research. Thus, we had access to the data collected among the first large sub-sample of participants, that is, a total of 66,394 women, and 50,784 men, participating at Time 1. Of these, 63,467 women and 50,775 men reported being in a relationship and responded to the relevant items. In addition, we use data from the Time 3 questionnaire for test-retest reliability (three months after T1, N=63,314 women) and from the Time 4 questionnaire for subsequent relationship dissolution (one year after T1, N=51,739). The reduction in N is partly due to the fact that part of the initial sample had not yet reached Time 4. Mean age of participants was 30.0 (sd=4.6) for women, and 32.7 (sd=5.3) for men.

The evaluation of the RS10 and the RS5 was based on both samples. Preliminary analyses were conducted on the full sample 1 and on the first subset of data available in sample 2 (N=2555). As the initial analyses provided evidence for the reliability and validity of the RS10 scale, the original set of items (see below) were retained for the subsequent data collection. For the sake of presentation, all results herein are based on the current full samples. The preliminary analyses yielded highly similar results, but the full sample provides more precise estimates and statistical power. Thus, the results section includes analyses of both the RS10 and RS5 for the full sample 2.

Measures and criteria

The Relationship Satisfaction (RS) scale

We aimed at constructing a unidimensional scale of 10 statements (RS10), based on items typical of those used in previously developed scales (7,9,10,13,14). Previous findings provided the opportunity for choosing optimal items across scales. Thus, the strategy involved choosing the best from the past. By reviewing the literature and published scales a set of item themes (e.g. 'satisfaction with marriage', 'problems in marriage', 'happy in my relationship') were identified, which appeared to reoccur across scales. An expert group of psychologists and psychometricians initially evaluated the items and formulations. Items were chosen that a) appeared in several scales (although with somewhat different wordings), b) that had high face validity as measures of global relationship satisfaction, c) yet were sufficiently distinct to avoid redundancy, d) that were judged to be relatively free of specific cultural content, and e) that focused on respondent evaluations rather than behaviors. Moreover, items were chosen that left open for the respondent to use her/his own subjective criteria to determine level of satisfaction. For example, the item "I am satisfied with the relationship to my partner" requires an evaluation of satisfaction, but this satisfaction is based on the subject's own values and criteria. Finally, we aimed at including both positively and negatively worded items, and a six-point response format was used, ranging from 'strongly disagree' to 'strongly agree'. The following items were included:

1. I have a close relationship with my spouse/partner
2. My partner and I have problems in our relationship
3. I am very happy with our relationship
4. My partner is generally understanding
5. I often consider ending our relationship
6. I am satisfied with my relationship with my partner
7. We frequently disagree on important decisions
8. I have been lucky in my choice of a partner
9. We agree on how children should be raised
10. I think my partner is satisfied with our relationship

Criteria and analyses

The following criteria were used for evaluating the quality of the scales:

Reliability was examined by calculating internal consistency (Cronbach's alpha) and test-retest correlations.

Structural validity involves confirming a hypothesized model of the underlying structure (20). The RS scale was hypothesized to have a unidimensional structure, with a single underlying latent factor accounting for item covariances, substantial factor loadings (i.e.

>.40) and a low degree of mutually redundant items (as indicated by residual correlations). Measurement invariance (21,22) across gender was also tested.

Convergent validity (23,24) was investigated by correlations with the established Quality Marriage Index (7) and correlations between spouses/partners. In support for the convergent validity of the scale we would expect a high correlation with an established measure of relationship satisfaction, and a moderate spousal agreement on relationship satisfaction (25).

Discriminant validity was investigated by correlations with general life satisfaction (26), and personality traits (27). As support for discriminant validity we expected neither life satisfaction nor personality traits to explain the larger part of the variance in relationship satisfaction.

Predictive validity was examined by means of relative risk and odds ratio for future relationship dissolution (divorce/break-up). A substantially increased risk of future dissolution, as predicted by the RS scale, would be further evidence of validity.

Other measures

Quality Marriage Index (QMI) (7) is a generic measure of marital quality including five items (e.g. "We have a good marriage") responded to on a 1-7 scale. The scale was included in sample 1.

The Satisfaction with Life Scale (SWLS) (26) measures global life satisfaction. The SWLS contains five items (e.g. "I am satisfied with my life") responded to on a 1-7 scale. The scale has been widely used in well-being research and has well-established psychometric properties (28,29). The scale was available in both samples.

Relationship dissolution (divorce/break-up) was measured by a single item in questionnaire 4, in the MoBa sample, collected approximately one year after questionnaire 1.

Big five personality traits were measured by scales from the International Personality Item Pool (IPIP-50) (27). The general personality traits of neuroticism, extraversion, agreeableness, conscientiousness and openness are measured by ten items each. The IPIP-50 was available for a subsample (N=8889) of the male respondents in MoBa.

Statistical analyses

In addition to the ordinary analyses conducted in SPSS, we used structural equation modeling in Mplus (21) to perform confirmatory factor analyses (CFA). Model fit was evaluated by the root-mean-square-error-of-approximation (RMSEA), and the comparative fit index (CFI). RMSEA values below .08 indicate acceptable fit, and values below .05 indicate good fit. CFI values above .90 (or preferably .95) also indicate good fit (30,31). Measurement invariance across gender was tested by comparing models with a) all parameters allowed to vary freely across gender, b) factor loadings constrained to be equal across gender, and finally c) both factor loadings and intercepts

constrained to be equal (21,32,33). Given the very large sample size, significance testing of $\Delta\chi^2$ would not be optimal. Rather, we compared the different fit indices, and adopted the recommendation that differences in CFI values between nested models should not exceed .01 (22).

RESULTS

We first report the results pertaining to the reduction of the RS10 into the short RS5. This reduction was done subsequent to the initial analyses providing evidence for the reliability and validity of the full RS10 scale (see below). However, here we outline the scale-reduction results first, thereafter we report psychometric findings for both scale versions.

For the purpose of reducing the full RS10 scale to a shorter RS5 scale we regressed the full scale (mean-score index) on the 10 individual items, by the stepwise method. This strategy optimizes the correlation between the full scale and the short scale, and has previously been applied successfully, for example for the Symptom Check List (SCL) (34). The first five items extracted were item number 2, 3, 6, 7 and 9 (see method section). Closer inspection of the data indicated that item 7 appeared difficult for a few respondents. This item involves a possible double negation (e.g. disagreeing on "We often disagree on important decisions"), potentially leading to errors in some cases. Therefore, we omitted this item from the short-list, included the next item, and retained the following items: 2, 3, 4, 6 and 9. These five items obtained R=.97 with the full RS10 scale (R^2=.94). In the following, this short RS5 was evaluated along the full RS10.

Descriptives and reliability

Table 1 shows descriptive statistics for the set of scales included here. Results are presented for the full sample 1, and for the larger sample 2 split by gender.

Table 1. Descriptives and reliability.

	Mean	SD	Alpha	Test-retest	% dis-satisfied	N
Sample 1						
RS10	4.95	0.95	.92		12.1	347
RS5	4.91	1.01	.88		14.4	347
SWLS	4.93	1.27	.92			347
QMI	5.92	1.28	.95			347
Sample 2 (women)						
RS10	5.30	0.66	.91	.77	4.6	63480
RS5	5.22	0.73	.85	.76	5.8	63467
SWLS	5.63	1.07	.89			66394
Sample 2 (men)						
RS10	5.32	0.58	.89		2.9	41963
RS5	5.28	0.63	.82		3.5	50775
SWLS	5.61	1.00	.86			50784
Extraversion	3.52	0.67	.89			8889
Neuroticism	3.63	0.65	.85			8889
Agreeableness	4.01	0.47	.80			8889
Conscientiousness	3.73	0.55	.80			8889
Openness	3.65	0.50	.75			8889

Table 2. Confirmatory factor analyses and invariance testing.

	χ^2	df	RMSEA	CFI	TLI
Sample 1					
RS10	76.82	35	0.059	0.97	0.97
RS5	12.56	5	0.066	0.99	0.98
Sample 2 (MoBa)					
RS10 (all)	9852.69	35	0.049	0.96	0.94
RS10 (women)	6622.03	35	0.054	0.96	0.94
RS10 (men)	3673.82	35	0.045	0.95	0.94
RS5 (all)	897.99	5	0.040	0.99	0.98
RS5 (women)	616.77	5	0.044	0.99	0.98
RS5 (men)	330.10	5	0.036	0.99	0.98
Invariance testing (MoBa)					
RS10 Baseline model	10188.37	70	0.050	0.95	0.94
RS10 Fac.loadings equal	10594.01	79	0.048	0.95	0.95
RS10 Fac.loadings & intercepts equal	11923.09	88	0.048	0.95	0.95
RS5 Baseline model	917.07	10	0.040	0.99	0.98
RS5 Fac.loadings equal	901.25	14	0.033	0.99	0.98
RS5 Fac.loadings & intercepts equal	1356.15	18	0.036	0.98	0.98

Table 3. Factor loadings from exploratory (EFA) and confirmatory (CFA) factor analyses, and item-total correlations, for the Relationship Satisfaction (RS10) scale. Total Moba sample, N=114528.

	Factor loadings		
Item	EFA	CFA	Item-total correlation
1. I have a close relationship with my spouse/partner	.84	.83	.81
2. My partner and I have problems in our relationship	.76	.72	.78
3. I am very happy with our relationship	.87	.87	.84
4. My partner is generally understanding	.76	.71	.76
5. I often consider ending our relationship	.72	.67	.71
6. I am satisfied with my relationship with my partner	.86	.85	.83
7. We frequently disagree on important decisions	.51	.43	.59
8. I have been lucky in my choice of a partner	.81	.78	.78
9. We agree on how children should be raised	.55	.49	.59
10. I think my partner is satisfied with our relationship	.81	.78	.79

Note. Items in *italics* are included in the short RS5 (i.e. items 2, 3, 4, 6 and 9).

Means and standard deviations, observed at Time 1, 18 weeks of gestation, were highly similar for women and men. Across gender and samples, internal reliability was high for both the RS10 and the RS5 scale. Test-retest correlations (from Time 1 to Time 3, across three months) were also relatively high. Table 1 also includes prevalence of 'dissatisfactory relationship', defined as a mean score below four (see below for cut-off details).

Validity

Structural validity was examined by factor analyses, based on the hypothesis that a single underlying latent factor could account for the observed covariance structure of the observed data. Exploratory factor analyses clearly suggested unidimensionality, with the Eigenvalues for the first two unrotated factors being 6.5 and 0.9 (sample 1) and 5.7 and 0.9 (sample 2). Further, we conducted confirmatory factor analyses with one latent factor. Table 2 shows fit of the measurement model. The single latent factor model yielded good fit to the data in both samples, for both the full and short scale.

Given that mutually redundant items would have contributed to multidimensionality and could have given a poor fit in the CFA, it seems warranted to consider the items to perform well.

Next, we tested the invariance of the measurement model across gender. For both the full and the short scale we first tested a baseline model in which all parameters were allowed to vary across gender. Second, we constrained the factor loadings to be equal across gender, and finally we constrained both the factor loadings and the intercepts to be equal (21,22). For both RS5 and RS10 the RMSEA value was lower for the invariance model (equal factor loadings and intercepts) than for the baseline model. The CFI and TLI values differed by less than 0.1 across the models, providing evidence of measurement invariance. Thus, support was found for the unidimensionality of the scales, and for invariance across gender. Table 3 shows factor-loadings from the exploratory and confirmatory factor analysis, as well as item-total correlations.

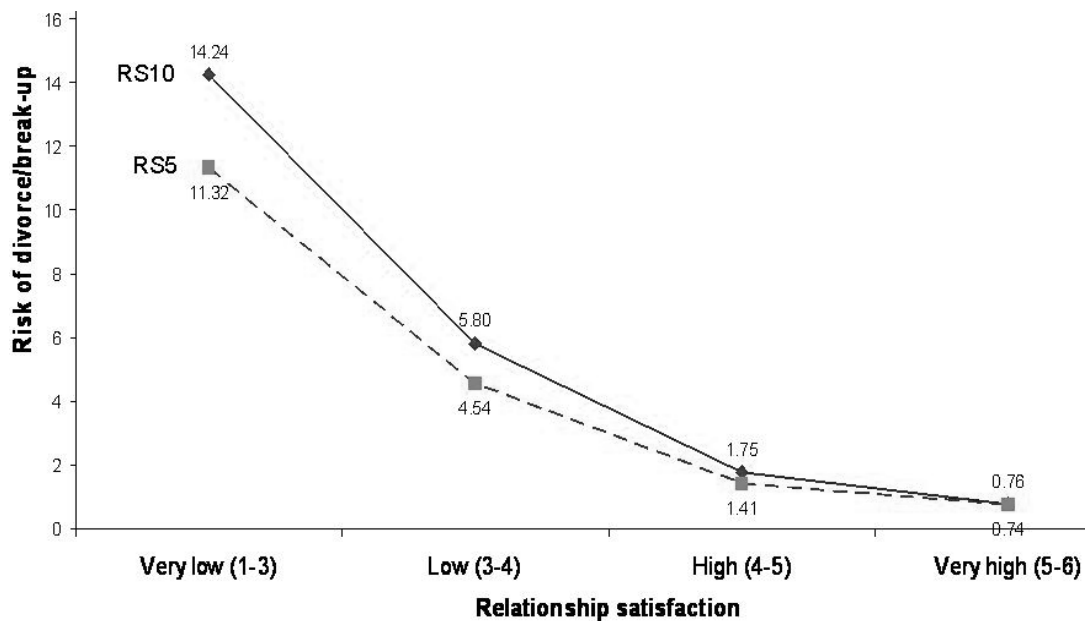Convergent and discriminant validity was investigated by correlations with other scales and across

Figure 1. Risk of relationship dissolution, within one year, for different levels of Relationship Satisfaction.

spouses. As can be seen in Table 4, the RS scales correlated highly with the Quality of Marriage Index, while agreement between spouses/partners was moderate. The correlation with general life satisfaction was low to moderate, and correlations with personality traits were relatively low. In addition to the correlation analyses we ran a multiple regression analysis with all Big Five traits as predictors of relationship satisfaction. Significant ($p<.01$) effects were found for neuroticism ($\beta=-.20$), agreeableness ($\beta=.18$) and conscientiousness ($\beta=.10$), yet personality traits explained only 12% of the total variance in relationship satisfaction.

Finally, to investigate predictive validity, the associations between the RS scale and relationship dissolution one year later was investigated. Figure 1 shows the risk of break-up depending on score on the RS scales. The scores were categorized into four groups, that is, scores in the following ranges, respectively: 1-3, 3-4, 4-5, 5-6 (with integer values assigned to the lower category, i.e. a score of 3.00 was assigned to the 1-3 group). The RS5 and RS10 were analyzed separately, with combined scores for women and men. As can be seen in Figure 1, a high (5-6) score on RS implies a low (<1%) risk of break-up within one year, while a low score (1-3) implies a relatively high (11-15%) risk of break-up. RS10 low scores were somewhat more predictive than RS5.

Although any cut-off point will be somewhat arbitrary, we suggest a mean score below '4' as a useful cut-off. Semantically, a score of four implies 'slightly agree' as the average response across all satisfaction items, and a score below 4 suggests a certain level of dissatisfaction. Further, this cut-off yields prevalence rates of 'dissatisfactory relationship' varying between 2.9% and 14.4% (see Table 1). Moreover, as seen in Figure 1, risk of future divorce increases substantially around this level. We conducted a set of logistic

regression analyses, with relationship dissolution as dependent variable, and a dichotomized RS (1-3.999 vs 4-6) as independent variable. The Odds Ratio (OR) for future break-up, given a low RS10 score, was 9.3 (CI: 7.8–11.2) for women's scores, and 8.0 (CI: 6.7–9.5) for men's scores. The corresponding results for RS5 were OR=7.8 (CI: 5.9–10.2) for women and OR=6.7 (CI: 5.1–8.6) for men. Further, this low-high dichotomized RS predicted dissolution vs continued relationship correctly for 93.8% of the women and 95.8% of the men.

DISCUSSION

We set out to examine the reliability and validity of the Relationship Satisfaction scale. The scale was developed within the context of the Norwegian Mother and Child Cohort Study and is based on several previously

Table 4. Correlations between the Relationship Satisfaction (RS) scales and external validity measures.

	RS10	RS5
Sample 1		
RS5	.98	
Quality Marriage Index	.92	.91
Satisfaction with Life Scale	.49	.49
Sample 2		
RS5	.97	
RS10-spouse	.57	.55
RS5-spouse	.55	.54
SWLS	.43	.41
Extraversion	.14	.14
Neuroticism	-.27	-.27
Agreeableness	.26	.26
Conscientiousness	.21	.21
Openness	.10	.10

$p<.001$ for all correlations

developed scales (10,13,14). Two versions of the scale were introduced, a full scale including ten items (RS10), and a short scale with five items (RS5). Results were provided from two population based samples including more than 100,000 participants in total. Sample 1 included a broad range of age groups and contained several measures that contributed to establishing the validity of the RS scale. Sample 2 included a large number of participants. This sample also included longitudinal data, providing the basis for investigating predictive validity, and identical measures for spouses, which allowed for investigating degree of agreement in perceptions of the relationship.

The scale was found to show good psychometric properties. For both scale versions, internal reliability and test-retest reliability was high. The structural validity was confirmed by factor analyses. A model of a single latent factor accounting for the covariance structure of the items fit the data well. The good fit also implies absence of nontrivial residual correlations, which would have been indicative of redundant items. Convergent validity was evidenced by a high correlation with the Quality of Marriage Index (7), and moderately high agreement between spouses/partners. The correlation with the general Satisfaction With Life Scale serves as an indicator of discriminant validity. That is, several studies have shown substantial relations between general life satisfaction and domain satisfaction (e.g. work-related, health-related, and relational satisfaction) (35,36). The moderate association between life satisfaction and relationship satisfaction found here accords with previous findings, and indicates that the two satisfaction measures contain unique variance rather than just reflect a general tendency to perceive various life domains positively or negatively.

Further, in accordance with previous studies relationship satisfaction was correlated with the Big Five personality traits (37). In particular, high scores on agreeableness and conscientiousness, and low scores on neuroticism, were predictive of high relationship satisfaction. However, none of the correlations exceeded .30, thus providing further support for the discriminant validity of the RS scale. That is, personality traits contribute to relationship satisfaction, yet, the major part of the variance of RS is unrelated to personality traits.

Finally, evidence of predictive validity was demonstrated by a strong association with subsequent relationship dissolution. A one year time span is limited, yet the respondents scoring in the lower part of the RS scale showed a substantially heightened risk of break-up in the coming year. As such, the scale has the potential to forecast future events. Another sub-study within the MoBa found that current levels of relationship satisfaction also predict divorce many years ahead (38) and predict later change in general life satisfaction (39). In particular, during the pre- and postnatal period, a good relationship appears to imply a relative increase in future life satisfaction.

It seems warranted to recommend the RS10 and RS5 as measures of general relationship satisfaction. The RS10 shows slightly superior results throughout the analyses (i.e. reliability, convergent and predictive validity). However, the differences were minor, and the RS5 performed best in some of the confirmatory factor analyses. Thus, both versions of the scale appear to capture the same underlying variance, and to perform well. Given sufficient space the RS10 can be recommended, but in studies with space limitations the RS5 provides essentially the same information and may be more suitable. Also, the RS10 can be seen as a pool of well-functioning items (with the possible exception of item 7). Researchers with special interests can choose other sub-sets of items than those included in the RS5 (for example, for couples not having children, item 9, addressing childrearing could be replaced with another item).

The scale was primarily developed for research purposes. However, we believe the scale might be a useful tool also in clinical settings. In marriage counseling and couples therapy the level of agreement between spouses, regarding their relationship satisfaction, might be an important object of discourse and exploration. Further, with the descriptive statistics provided by two population based samples, and the suggested cut-off level of 'dissatisfactory relationship', the scale has a potential for identification of at-risk couples in clinical contexts.

Some limitations of the study should be noted. Despite the strengths of being population based, and having a large sample, the MoBa study has a limited response-rate. Thus, in particular the descriptive statistics (mean and prevalence) might be biased as the participants represent a relatively healthy group. However, the nature of couple relationships involves the basic option of ending the relationship if it is perceived as dissatisfactory over time. Thus, it would be expected that most people in a relationship should be more satisfied than dissatisfied, otherwise they would not be there. We believe the MoBa sample consists of relatively well-adapted and happy participants, leading to high scores on satisfaction measures. Yet, there is no reason to consider such a result a threat to the findings of factor structure and associations with other variables. Due to the generally high scores, and in line with other similar scales (7,14) the RS scales were negatively skewed. For purposes requiring normalized variables a log-transformation (of reversed scale) is recommended.

In conclusion, the Relationship Satisfaction scale shows excellent psychometric properties. Hopefully the scale will be found useful for research and clinical practice. We encourage further studies addressing questions about the causes, correlates and consequences of different levels of relationship satisfaction to consider using it. Future research should also address issues of stability and change in relationship satisfaction, and long term predictive validity.

REFERENCES

1. Stack S, Eshleman JR. Marital status and happiness: A 17-nation study. *J Marriage Fam* 1998;**60**:527-36.
2. Bradbury TN, Fincham FD, Beach SRH. Research on the nature and determinants of marital satisfaction: A decade in review. *J Marriage Fam* 2000;**62**:964-80.
3. Goldstein JR. The leveling of divorce in the United States. *Demography* 1999;**36**(3):409-14.
4. Amato PR, Booth A. *A generation at risk: Growing up in an era of family upheaval*. Cambridge, MA: Harvard University Press, 1997.
5. Cherlin AJ. American marriage in the early twenty-first century. *Future of Children* 2005;**15**(2):33-55.
6. Lester D. Trends in divorce and marriage around the world. *J Divorce Remarriage* 1996;**25**(1-2):169-71.
7. Norton R. Measuring marital quality: A critical look at the dependent variable. *J Marriage Fam* 1983;**45**:141-51.
8. Spanier G. Measuring dyadic adjustment: New scales for assessing the quality of marriage and similar dyads. *J Marriage Fam* 1976;**38**:15-28.
9. Snyder DK. *Marital Satisfaction Inventory–Revised (MSI-R) manual*. Los Angeles: Western Psychological Services, 1997.
10. Schumm WR, Paffbergen LA, Hatch RC, Obiorah FC, Copeland JM, Meens LD, et al. Concurrent and discriminant validity of the Kansas Marital Satisfaction Scale. *J Marriage Fam* 1986;**48**(2):381-7.
11. Jones WH, Adams JM, Monroe PR, Berry JO. A psychometric exploration of marital satisfaction and commitment. *J Sol Behavior Per*1995;**10**(4):923-32.
12. Stanley SM, Markman HJ. Assessing Commitment in Personal Relationships. *J Marriage Fam* 1992;**54**(3): 595-608.
13. Henrick SS. A generic measure of relationship satisfaction. *J Marriage Fam* 1988;**50**:93-8.
14. Blum JS, Mehrabian A. Personality and temperament correlates of marital satisfaction. *J Pers* 1999;**67**(1):93-125.
15. Funk JL, Rogge RD. Testing the ruler with item response theory: Increasing precision of measurement for relationship satisfaction with the couples satisfaction index. *J Fam Psychol* 2007;**21**(4):572-83.
16. Péloquin K, LaFontaine M-F. Measuring empathy in couples: Validity and reliability of the Interpersonal Reactivity Index for couples. *J Pers Assess* 2010;**92**(2):146-57.
17. Herrington RL, Mitchell AE, Castellani AM, Joseph JI, Snyder DK, Gleaves DH. Assessing disharmony and disaffection in intimate relationships: Revision of the marital satisfaction inventory factor scales. *Psychol Assess* 2008;**20**(4):341-50.
18. Magnus P, Irgens LM, Haug K, Nystad W, Skjærven R, Stoltenberg C. Cohort profile: The Norwegian Mother and Child Cohort Study (MoBa). *Int J Epidemiol* 2006;**35**(5):1146-50.
19. Nilsen RM, Vollset SE, Gjessing HK, Skjærven R, Melve KK, Schreuder P, et al. Self-selection and bias in a large prospective pregnancy cohort in Norway. *Paediatr Perinat Epidemiol* 2009;**23**(6):597-608.
20. John OP, Benet-Martinez V. Measurement: Reliability, construct validation, and scale construction. In: Reis HT, Judd CM, editors. Handbook of research methods in social and personality psychology. Cambridge: Cambridge University Press, 2000.
21. Muthén LK, Muthén BO. *Mplus User's Guide*, 4th edn. Los Angeles, CA: Muthén & Muthén, 2006.
22. Cheung GW, Rensvold RB. Evaluating goodness-of-fit indexes for testing measurement invariance. *Struct Eq Mod* 2002;**9**(2):233-55.
23. Carmines EG, Zeller RA. *Reliability and Validity Assessment*. London: Sage University Papers, 1980.
24. Brown TA. *Confirmatory factor analysis for applied research*. New York: The Guilford press, 2006.
25. Karney BR, Bradbury TN. The longitudinal course of marital quality and stability – a review of theory, method, and research. *Psych Bull* 1995;**118**(1):3-34.
26. Diener E, Emmons RA, Larsen RJ, Griffin S. The Satisfaction With Life Scale. *J Pers Assess* 1985;**49**(1):71-5.
27. Gow AJ, Whiteman MC, Pattie A, Deary IJ. Goldberg's 'IPIP' Big-Five factor markers: Internal consistency and concurrent validation in Scotland. *Pers Ind Diff* 2005;**39**(2):317-29.
28. Diener E, Oishi S, Lucas RE. Personality, culture, and subjective well-being: emotional and cognitive evaluations of life. *Ann Rev Psychol* 2003;**54**:403-25.
29. Pavot W, Diener E, Colvin CR, Sandvik E. Further validation of the Satisfaction with Life Scale: evidence for the cross-method convergence of well-being measures. *J Pers Assess* 1991;**57**(1):149-61.

30. Hoyle RH. *Structural equation modeling: concepts, issues and applications*. London: Sage, 1995.
31. Bollen KA, Curran PJ. *Latent Curve Models: A Structural Equation Perspective*. New York: Wiley-Interscience, 2006.
32. Rensvold RB, Cheung GW. Testing measurement models for factorial invariance: A Systematic approach. *Edu Psychol Measure* 1998;**58**(6):1017-34.
33. Vandenburg RJ, Lance CE. A review and synthesis of the measurement invariance literature: Suggestions, practices, and recomendations for organizational research. *Org Res Met* 2000;**3**:4-70.
34. Tambs K, Moum T. How well can a few questionnaire items indicate anxiety and depression? *Acta Psychiatr Scand* 1993;**87**(5):364-7.
35. Diener E, Diener M. Cross-cultural correlates of life satisfaction and self-esteem. Culture and well-being: The collected works of Ed Diener. New York, NY: Springer Science + Business Media, 2009: 71-91.
36. Oishi S, Diener E, Lucas RE, Suh EM. Cross-cultural variations in predictors of life satisfaction: Perspectives from needs and values. Culture and well-being: The collected works of Ed Diener. New York, NY: Springer Science + Business Media; 2009: 109-27.
37. White JK, Hendrick SS, Hendrick C. Big five personality variables and relationship constructs. *Pers Ind Diff* 2004;**37**:1519-30.
38. Rosand GMB, Slinning K, Røysamb E, Tambs K. Relationship dissatisfaction and other risk factors for future relationship dissolution: a population-based study of 18,523 couples. *Soc Psych Psych Epid* 2014;**49**(1):109-19.
39. Dyrdal GM, Røysamb E, Nes RB, Vitterso J. Can a happy relationship predict a happy life? A population-based study of maternal well-being during the life transition of pregnancy, infancy, and toddlerhood. *J Happiness Stud* 2011;**12**(6):947-62.