PUBLISHED CLINICAL PSYCHOLOGY & PSYCHOTHERAPY

<https://doi.org/10.1002/cpp.2420>

**Are Measures of Marital Satisfaction Valid for Women with Depressive Symptoms?
The Examination of Factor Structure and Measurement Invariance of the Couple
Satisfaction Index-4 across depression levels in Portuguese Women**

Diogo Lamela, Bárbara Figueiredo, Ana Morais, Paula Matos & Inês Jongenelen

Abstract

Marital satisfaction (MS) is a key indicator of mental and physical health. Factor structure of MS measures in individuals with clinical levels of depression as well as their measurement invariance across groups with different levels of depression symptoms were not yet explored. The lack of evidence of measurement invariance might compromise valid comparisons between individuals with high and low levels of depression symptoms in MS. This study examined the factor structure of the Couple Satisfaction Index-4 (CSI-4) among women with clinical levels of depression, tested the CSI-4 measurement invariance across depression levels groups, and investigated CSI-4 convergent and divergent validity. Participants were 891 heterosexual married/cohabiting women who were assigned into one of two groups based on assessment of their levels of depression symptoms. Participants completed the CSI-4 and self-reported measures used to examine convergent and divergent validity. Support was found for the factor structure of the CSI-4 for the total sample and both high levels of depression symptoms and low levels of depression symptoms groups. Subsequent multigroup confirmatory factor analyses supported the measurement invariance of the CSI-4 across depression levels groups. The CSI-4 revealed excellent values of internal consistency and convergent and divergent validity. Our findings suggest that CSI-4 produces comparable response patterns across depression groups and thus meaningful comparisons between groups can be performed.

Keywords: marital satisfaction; depression; measurement invariance; couple; factor structure

Marital satisfaction (MS) is conceptualized as a key dimension on the understanding of the bidirectional effects between marital functioning and depression symptoms (Rehman, Gollan, & Mortimer, 2008; Trombello, Post, & Smith, 2018). Research has documented strong associations between poor MS and the onset, relapse, chronicity, and prognosis of depression symptoms (van Loo, Aggen, Gardner, & Kendler, 2017). Prior empirical work has also examined the moderating effect of depression on the associations between women's MS and clinical outcomes. Results of this body of research have demonstrated that depressed women with low MS showed higher levels of ineffective parenting, poorer health outcomes, and lower response to couple-based clinical interventions than marital dissatisfied non-depressed women (Baucom et al., 2017; Kouros, Papp, Goeke-Morey, & Cummings, 2014; Robles, Slatcher, Trombello, & McGinn, 2014).

Although MS measures have systematically been used to compare depressed and non-depressed women in the assessment of their romantic relationship, to our knowledge no study examined the factor structure of such measures with individuals with clinical levels of depression as well as their measurement invariance across depressed and non-depressed groups. This is a critical methodological issue since researchers have been assuming that the factor structure and psychometric properties of these measures would be similar in both non-clinical and clinical populations. However, it is plausible that depression-related cognitive and emotional impairments may bias depressed individuals' interpretation of item content and also their expectations and reporting standards for MS (Everaert, Duyck, & Koster, 2014; Joormann & Vanderlind, 2014; Smith & Peterson, 2008). Such systematic reporting differences between depressed and non-depressed individuals could elicit significant over- or underestimates of MS and consequently to distinct factor structures. Thus, without establishing scales' measurement invariance, meaningful comparisons between depressed and

non-depressed groups in MS are largely compromised due the absence of psychometric evidence regarding whether a set of items will represent similarly the targeted latent construct in both groups (Adolf, Schuurman, Borkenau, Borsboom, & Dolan, 2014; Cheung & Rensvold, 2002). As a result, researchers were not able to guarantee whether potential differences in MS between depressed and non-depressed groups and subsequent associations with other variables would reflect theory-inferred underlying psychological processes or instead the lack of measurement accuracy (Adolf et al., 2014). Though, in order to perform valid comparisons between depression-based groups in MS, multi-group confirmatory factor analysis should previously be conducted to establish whether a given measure produces comparable response patterns across samples (Brown, Harris, O'Quin, & Lane, 2017; Fried et al., 2016). In case that statistical equivalence in responding is obtained, then measure score comparisons become possible and samples can be assumed to be from the same population.

To overcome this methodological weakness, the current study sought to examine the factor structure and the measurement invariance of a widely used measure of MS across depressed and non-depressed groups. The CSI-4 was developed to assess individuals' global evaluations of their romantic relationship (Funk & Rogge, 2007) and its major strength is the highest levels of precision and power than traditional measures of MS (e.g., Marital Adjustment Test and Dyadic Adjustment Inventory [DAS; Spanier, 1976]). The CSI-4 shows a wide range of applications, being a common measure in the assessment of marital satisfaction in depressed individuals (Whitton & Kuryluk, 2014) and in the evaluation of couple therapy efficacy (Baucom et al., 2017). It is also used in studies with a wide range of samples, from individuals exposed to intimate violence (Visschers, Jaspaert, & Vervaeke, 2017) to participants diagnosed with a medical condition (Saiki & Meize-Grochowski, 2017).

The original psychometric study found a unidimensional factor structure of the CSI-4 (Funk & Rogge, 2007), which was also confirmed in international samples (Qadir, Khalid,

Haqqani, Zill-e-Huma, & Medhin, 2013). In addition, the American version of the CSI-4 showed excellent levels of internal consistency and convergent validity (e.g., r of .87 with DAS) (Funk & Rogge, 2007). The CSI has 3 versions (36, 16, and 4 items) highly correlated to each other. However, similarly to other measures, the evaluation of the psychometric properties of the CSI-4 for depressed individuals along with factor invariance across different depression levels groups were not yet investigated.

In this study, we examined (a) factor structure by testing the factor structure of the Portuguese version of the CSI-4, (b) measurement invariance across depression levels groups, (c) reliability of the scale (internal consistency), (d) convergent validity of the CSI-4 with a measure of marital instability, and (e) divergent validity with a measure of social desirability.

Methods and materials

Participants and procedures

The participants were a subset of 891 women with a current romantic relationship from a larger online community-based study designed to collect information regarding mental health, sexual behavior, and family functioning in Portuguese women. Demographic data are summarized in Table 1. To be eligible in this research, respondents had to be at least 18 years old and be resident Portugal (for a detailed description about the research protocol, recruitment strategy and data cleaning procedures see (Lamela, Jongenelen, Morais, & Figueiredo, 2017). In the current study, only heterosexual married or cohabiting women were included in the further analyses.

All participants were firstly screened for depression symptoms using the Portuguese version of the Patient Health Questionnaire-9 (PHQ-9; (Kroenke, Spitzer, & Williams, 2001). Women were assigned into the low levels of depression symptoms group or the high levels of depression symptoms group based on PHQ-9 answers. Participants were included in the high

levels of depression group ($n = 216$) if they fulfilled criteria for at least subthreshold symptoms of depression according to DSM-V criteria for Minor Depression as indicated by at least two criteria of the Major Depression symptoms present during the same 2-week period, and at least one of the symptoms of either depressed mood or markedly diminished interest or pleasure in activities, as assessed by the PHQ-9. No differences were found between groups in sociodemographic variables (Table 1). Ethical approval of this study was granted by the by the Scientific Committee of the Faculty of Psychology, Education and Sports of the Lusófona University of Porto.

Measures

The CSI-4 measures individuals' global evaluations of their romantic relationship (Funk & Rogge, 2007). It includes four items that are rated on a 7-point Likert scale for the first item (0 = *extremely unhappy*; 6 = *perfect*) and on a 6-point Likert scale for the other three items (0 = *not at all*; 5 = *completely*). Items' scores are summed so that higher total scores reflect higher levels of global marital satisfaction (total score ranging from 0 to 21). The Portuguese version of the CSI was developed based on back translation procedure (AUTHORS MASKED REFERENCE).

Marital Instability Index-Adapted (MII-Ad; (Booth, Johnson, & Edwards, 1983) assesses the frequency in last 3 months of thoughts about relationship difficulties, talking with friends about marital dissolution, desire of marital dissolution, and thoughts about the post-dissolution life. The MII-Ad comprises 4 items rated in a 4-point Likert scale (0 'Never' to 3 'A lot'). Higher scores reflect higher marital instability. The Portuguese version of the MAI-Ad showed excellent psychometric properties (AUTHORS MASKED REFERENCE). Cronbach's alpha for the current sample was .89.

The Marlowe-Crowne Social Desirability Scale-SF (MCSDS-SF;(Ballard, 1992)) is a forced-choice (true/false) 5-item self-report measure designed to assess the tendency to

present oneself as in a more favorable way than actually is. Higher scores indicate stronger tendency to portray oneself in a socially desirable light. The Portuguese version of this scale showed good psychometric properties (Pechorro, Vieira, Poiares, & Marôco, 2012).

Cronbach's alpha for the current sample was .81.

The PHQ-9 measures the frequency of depressive symptoms experienced in the past two weeks. This 9-item scale assesses all DSM-V symptoms of Major Depression. Each item is answered along a 4-point Likert scale (from 0 'not at all' to 3 'nearly every day'). Higher scores reflect higher severity of depression symptoms. The PHQ-9 may also be employed as a diagnostic measure (Kroenke et al., 2001). As a diagnostic tool, Major Depression is diagnosed if ≥ 5 of the 9 symptoms have been present at least more than half the days in the past 2 weeks and one of these symptoms is either depressed mood or anhedonia. Minor depression is diagnosed if 2 a 4 symptoms have been present at least more than half the days in the past two weeks and one of the symptoms is either depressed mood or anhedonia. Its internal consistency in this sample was .89.

Statistical strategy

First, a Confirmatory Factor Analyses (CFA) (with full information maximum likelihood estimation) was conducted to test the factor structure of the CSI-4. CFI, NNFI, RMSEA, and SRMR were used as model fit indexes. As standard rules of thumb, values of $\geq .95$ in CFI and NNFI, and values of ≤ 0.05 in RMSEA and SRMR suggest good model fit (Kline, 2015). Three measurement invariance tests were conducted between depression levels groups: configural invariance (invariance in factor structure), metric invariance (invariance in items' factor loadings), and scalar invariance (invariance of items' intercepts) (Chen, 2007; Cheung & Rensvold, 2002). A ΔCFI of $\leq .01$ between a baseline model and the resulting model was used as indicator of measurement invariance (Cheung & Rensvold, 2002).

Cronbach's alphas were calculated to examine internal consistency, with values of $\geq .90$ suggesting excellent internal consistency (Nunnally & Bernstein, 1994). Convergent and divergent validity were examined conducting Pearson correlations. The lavaan package in R statistical software (Rosseel, 2012) was used to conduct CFA and MGCFA. The remaining analyses were performed using IBM SPSS 23.

Results

Factor structure and measurement invariance

Based on previous findings about the dimensionality of the CSI-4, a CFA was conducted to test the established unidimensional model, with four items defined as indicators for one latent factor (CSI). Table 2 shows the results of the CFA for the total sample in the first row. The 1-factor model showed a very good fit to the data, confirming the one-factor solution as reported in the American study. Factor loadings ranged from .65 to .96. Next, to test the factor invariance across depression levels groups, the same model was computed for the two groups. Table 2 shows the results of these CFAs. The model fit measures supported the one-factor solution for both high levels of depression symptoms and low levels of depression symptoms groups.

The results of the sequence of gradually more restrictive tests of measurement invariance are reported in Table 2. First, the tested model revealed a good fit to data when only the factor structure was constrained to be equal for both groups (configural invariance). Next, after constraining factor loadings to be equal across groups, results revealed that the model continued to demonstrate good fit ($\Delta CFI \leq .01$), providing support for metric invariance across depression levels groups. Finally, intercepts were constrained to be equal across groups. Results indicated that model continued to show good fit ($\Delta CFI \leq .01$), so intercepts may also be considered equivalent across groups (scalar invariance).

Internal consistency and convergent and divergent validity

The mean of the CSI-4 total scores is presented in Table 1. Excellent values of internal consistency were found, with α of .91, .92, .90 for total sample, no-depression group, and depression group, respectively. For the total sample, CSI-4 was strongly and negatively associated with MII-Ad ($r = -.84, p < .001$) and weakly and negatively associated with MCSDS-SF ($r = -.07, p < .05$). For the high levels of depression symptoms group, CSI-4 was strongly and negatively associated with MII-Ad ($r = -.79, p < .001$) and no significant association was found between CSI-4 and MCSDS-SF ($r = -.05, p = .45$). Finally, for the group with low levels of depression symptoms, CSI-4 was strongly and negatively associated with MII-Ad ($r = -.86, p < .001$) and no significant association were found between CSI-4 and MCSDS-SF ($r = -.07, p = .07$).

Discussion

Findings replicated the original 1-factor structure for total sample and both non-depressed and depressed groups. Consistently with prior empirical work (Qadir et al., 2013), our findings contribute to literature by showing the robustness of CSI-4 factor structure in the assessment of MS in distinct cultural backgrounds. This is particularly relevant since the factor models of some prior American measures showed poorer fit or failed to be replicated in other cultures (Lim & Ivey, 2000), suggesting the cultural-specific nature of items' content of those measures, which constrains their application in other countries and cross-cultural comparisons. Thus, the support found in our study for the factor structure may indicate that CSI-4 items not only can accurately capture MS but also are less sensitive to potential cultural biases in the operationalization of MS.

Good CSI-4 measurement invariance was also found across depression groups. Thus, our findings demonstrated that the factor loadings and factor variances and covariances were equivalent across depression levels groups in Portuguese women. Therefore, it may be recommend administering the Portuguese CSI-4 for women with any depression levels. This

finding assumes particular clinical and research utility since no previous study had investigated the measurement invariance of MS scales across groups with different levels depression symptoms. By demonstrating the statistical equivalence of the factor structure, our results highlight that CSI-4 produces comparable response patterns across depression levels groups and thus valid interpretations of total scores along with explainable and meaningful comparisons between groups can be performed. Our findings suggest that interpretation of item content and their expectations and reporting standards for MS did not differ between individuals with high levels of depression and individuals with low levels of depression, in contrast with the depression–distortion hypothesis (Richters & Pellegrini, 2010). In particular, this hypothesis postulates that dysphoric emotions associated with depression activate a negative perceptual bias in the individuals' ratings of a target construct (e.g., children's behavior problems, MS) that, in turn, leads to overreporting of difficulties in that variable of interest (Daryanani et al., 2015). Despite depression-distortion bias was documented in mothers' reports of child behavior problems (Daryanani et al., 2015) no empirical evidence was however found of a similar pattern of findings in MS in our study.

Results also revealed that the internal consistency coefficients of CSI-4 total score were excellent for the total sample and both groups, being comparable to those obtained in Funk and Rogge (Funk & Rogge, 2007). As expected, CSI-4 revealed a strong negative correlation with marital instability (MII-Ad), supporting with convergent validity of the scale. This finding is consistent with prior clinical research that has suggested that no gains in MS during couple therapy as reported by women were associated with later marital instability and subsequent divorce risk, when compared with women who reported increases in MS during therapy (Christensen, Atkins, Baucom, & Yi, 2010).

In addition, a very weak association (for total sample) or no association (both groups) was found between CSI-4 and social desirability. As in previous findings (Visschers et al.,

2017), our results suggest that CSI-4 scores are not significantly correlated with a response bias caused by individuals' tendency to present themselves in the most positive manner relative to predominant social norms. This finding not only supported the CSI-4 divergent validity (i.e., very weak/no correlation with a theoretically unrelated construct) but also suggested that CSI-4 validity may be not highly threaten by social desirability in contrast with other measures that failed to demonstrate that MS reports were not significantly influenced by social desirability (Schumm et al., 1986).

Two limitations warrant discussion. First, depression symptoms were only assessed by a self-report measure. Despite of the diagnostic validity of the PHQ-9 is firmly established, the measure does not exclude physical causes of depression, normal bereavement, and manic symptoms. The use of a diagnostic interview would provide additional accuracy in the clinical diagnosis of depression disorders. Second, the current research was only conducted with heterosexual married/cohabiting women. While this relative sample homogeneity augments statistical confidence in the associations found, this limited variability restrains the generalization of these findings to men, same-sex relationships, and other marital status. Thus, future research should expand these findings by using clinical interview for depression assessment, by conducting multigroup CFA across gender and sexual orientation, and by examining measurement invariance of other widely used measures of marital satisfaction.

Key Practitioner Message

- Measurement invariance of MS measures across depression levels is unexplored.
- Valid comparisons across groups with different levels depression in MS is compromised.
- The factor structure of the CSI-4 for women with high levels of depressive symptoms was supported.
- Measurement invariance of the CSI-4 across depression levels was found.

- Reliability, convergent and divergent validity of the Portuguese CSI-4 was also supported.

Funding information

This study did not receive any funding grant.

Conflicts of interest

None declared.

References

- Adolf, J., Schuurman, N. K., Borkenau, P., Borsboom, D., & Dolan, C. V. (2014). Measurement invariance within and between individuals: a distinct problem in testing the equivalence of intra- and inter-individual model structures. *Frontiers in Psychology*, 5, 883. <https://doi.org/10.3389/fpsyg.2014.00883>
- Ballard, R. (1992). Short forms of the Marlow-Crowne Social Desirability Scale. *Psychological Reports*, 71(3f), 1155–1160. <https://doi.org/10.2466/pr0.1992.71.3f.1155>
- Baucom, D. H., Fischer, M. S., Worrell, M., Corrie, S., Belus, J. M., Molyva, E., & Boeding, S. E. (2017). Couple-based Intervention for Depression: An Effectiveness Study in the National Health Service in England. *Family Process*. <https://doi.org/10.1111/famp.12332>
- Booth, A., Johnson, D., & Edwards, J. N. (1983). Measuring Marital Instability. *Journal of Marriage and the Family*, 45(2), 387. <https://doi.org/10.2307/351516>
- Brown, G. T. L., Harris, L. R., O'Quin, C., & Lane, K. E. (2017). Using multi-group confirmatory factor analysis to evaluate cross-cultural research: identifying and understanding non-invariance. *International Journal of Research & Method in Education*, 40(1), 66–90. <https://doi.org/10.1080/1743727X.2015.1070823>
- Chen, F. F. (2007). Sensitivity of Goodness of Fit Indexes to Lack of Measurement Invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 14(3), 464–504.

<https://doi.org/10.1080/10705510701301834>

- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating Goodness-of-Fit Indexes for Testing Measurement Invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 9(2), 233–255. https://doi.org/10.1207/S15328007SEM0902_5
- Christensen, A., Atkins, D. C., Baucom, B., & Yi, J. (2010). Marital status and satisfaction five years following a randomized clinical trial comparing traditional versus integrative behavioral couple therapy. *Journal of Consulting and Clinical Psychology*, 78(2), 225–235. <https://doi.org/10.1037/a0018132>
- Daryanani, I., Hamilton, J. L., Shapero, B. G., Burke, T. A., Abramson, L. Y., & Alloy, L. B. (2015). Differential Reporting of Adolescent Stress as a Function of Maternal Depression History. *Cognitive Therapy and Research*, 39(2), 110–119. <https://doi.org/10.1007/s10608-014-9654-4>
- Everaert, J., Duyck, W., & Koster, E. H. W. (2014). Attention, interpretation, and memory biases in subclinical depression: A proof-of-principle test of the combined cognitive biases hypothesis. *Emotion*, 14(2), 331–340. <https://doi.org/10.1037/a0035250>
- Fried, E. I., van Borkulo, C. D., Epskamp, S., Schoevers, R. A., Tuerlinckx, F., & Borsboom, D. (2016). Measuring depression over time . . . Or not? Lack of unidimensionality and longitudinal measurement invariance in four common rating scales of depression. *Psychological Assessment*, 28(11), 1354–1367. <https://doi.org/10.1037/pas0000275>
- 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>
- Joormann, J., & Vanderlind, W. M. (2014). Emotion Regulation in Depression. *Clinical Psychological Science*, 2(4), 402–421. <https://doi.org/10.1177/2167702614536163>

- Kline, R. (2015). *Principles and Practice of Structural Equation Modeling*. New York: Guilford.
- Kouros, C. D., Papp, L. M., Goeke-Morey, M. C., & Cummings, E. M. (2014). Spillover between marital quality and parent–child relationship quality: Parental depressive symptoms as moderators. *Journal of Family Psychology*, 28(3), 315–325.
<https://doi.org/10.1037/a0036804>
- Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine*, 16(9), 606–613.
<https://doi.org/10.1046/j.1525-1497.2001.016009606.x>
- Lamela, D., Jongenelen, I., Morais, A., & Figueiredo, B. (2017). Cognitive-affective depression and somatic symptoms clusters are differentially associated with maternal parenting and coparenting. *Journal of Affective Disorders*, 219, 37–48.
<https://doi.org/10.1016/j.jad.2017.05.006>
- Lim, B. K., & Ivey, D. (2000). The Assessment of Marital Adjustment with Chinese Populations: A Study of the Psychometric Properties of the Dyadic Adjustment Scale. *Contemporary Family Therapy*, 22(4), 453–465.
<https://doi.org/10.1023/A:1007801018478>
- Nunnally, J., & Bernstein, I. (1994). *Psychometric Theory*. New York: McGraw-Hill.
- Pechorro, P., Vieira, R. X., Poiares, C., & Marôco, J. (2012). Contributos para a Validação duma Versão Curta da Escala de Desejabilidade Social de Marlowe-Crowne com Adolescentes Portugueses. *Arquivos de Medicina*, 26(3), 103–108.
- Qadir, F., Khalid, A., Haqqani, S., Zill-e-Huma, & Medhin, G. (2013). The association of marital relationship and perceived social support with mental health of women in Pakistan. *BMC Public Health* 2013 13:1, 13(1), 1150. <https://doi.org/10.1186/1471-2458-13-1150>

- Rehman, U. S., Gollan, J., & Mortimer, A. R. (2008). The marital context of depression: Research, limitations, and new directions. *Clinical Psychology Review, 28*(2), 179–198. <https://doi.org/10.1016/J.CPR.2007.04.007>
- Robles, T. F., Slatcher, R. B., Trombello, J. M., & McGinn, M. M. (2014). Marital quality and health: A meta-analytic review. *Psychological Bulletin, 140*(1), 140–187. <https://doi.org/10.1037/a0031859>
- Saiki, L., & Meize-Grochowski, R. (2017). Urinary Incontinence and Psychosocial Factors Associated With Intimate Relationship Satisfaction Among Midlife Women. *Journal of Obstetric, Gynecologic & Neonatal Nursing, 46*(4), 555–566. <https://doi.org/10.1016/J.JOGN.2017.02.003>
- Smith, D. A., & Peterson, K. M. (2008). Overperception of Spousal Criticism in Dysphoria and Marital Discord. *Behavior Therapy, 39*(3), 300–312. <https://doi.org/10.1016/J.BETH.2007.09.002>
- Trombello, J. M., Post, K. M., & Smith, D. A. (2018). Depressive Symptoms, Criticism, and Counter-Criticism in Marital Interactions. *Family Process*. <https://doi.org/10.1111/famp.12349>
- van Loo, H. M., Aggen, S. H., Gardner, C. O., & Kendler, K. S. (2017). Sex similarities and differences in risk factors for recurrence of major depression. *Psychological Medicine, 1*–9. <https://doi.org/10.1017/S0033291717003178>
- Vischers, 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>
- Whitton, S. W., & Kuryluk, A. D. (2014). Associations between relationship quality and depressive symptoms in same-sex couples. *Journal of Family Psychology, 28*(4), 571–576. <https://doi.org/10.1037/fam0000011>

Table 1.

Sample Description and Group Differences in Sociodemographic Variables and Measures

Total sample (<i>N</i> = 891)	Low levels of depression group	High levels of Depression group (<i>n</i> = 216)	<i>t</i> (2, 889)
-----------------------------------	--------------------------------------	---	-------------------

	(N = 665)						
	M	SD	M	SD	M	SD	
Age	39.8	6.3	39.7	6.2	39.3	6.5	1.04
Years of education	16.3	5.7	16.5	6.1	15.8	4.2	1.50
Income (€/month)	1134.6	703.2	1176.6	707.3	1081.3	667.0	1.73
Relationship length (years)	13.4	6.8	13.4	6.7	12.5	6.2	1.74
Number of children	1.68	0.79	1.67	0.77	1.68	0.86	-0.126
CSI-4	12.2	5.4	12.6	5.5	11.1	4.9	3.26***
MII-Ad	3.10	3.57	2.96	3.62	3.45	3.43	-1.62
MCSDS-SF	8.22	1.27	8.12	1.25	8.50	1.27	-3.97***
PHQ-9	6.23	5.16	3.92	2.33	13.04	5.21	-35.73***

*** $p < .001$

Table 2.

Factor Loadings and Fit Indices for the Estimated 1-factor Model of the CSI-4 in the Total Sample and in Depression Levels Groups

	Factor Loading of CSI-4 Items ^{a,b}				<i>df</i> , χ^2	CFI	NNFI	RMSEA	SRMR	Δ CFI
	Item 1	Item 2	Item 3	Item 4						
Total sample (<i>N</i> = 891)	.65	.95	.96	.97	2; 3.37	0.99	0.99	0.028	0.004	
Depression groups										
High levels of depression (<i>n</i> = 226)	.58	.93	.98	.95	2; 3.09	0.98	0.98	0.045	0.010	
Low levels of depression (<i>n</i> = 665)	.66	.96	.97	.97	2; 3.55	0.99	0.99	0.034	0.007	
Multigroup analysis										
Configural invariance	–	–	–	–	4; 6.64	0.99	0.99	0.039	0.005	–
Metric invariance	–	–	–	–	7; 10.80	0.99	0.99	0.035	0.011	0.000
Scalar invariance	–	–	–	–	11; 34.94	0.99	0.99	0.051	0.025	0.003

^a CSI-4 items: 'Please indicate the degree of happiness, all things considered, of your relationship' (item 1), 'I have a warm and comfortable relationship with my partner' (item 2), 'How rewarding is your relationship with your partner?' (item 3), and 'In general, how satisfied are you with your relationship?' (item 4). ^b In all cases, $p < .001$.