Iranian Evolutionary Educational Psychology Journal
IEEPJ
University of Hormozgan
Online ISSN: 2588 - 4395
Homepage: https://ieepj.hormozgan.ac.ir

Predicting Marital Relationship Quality Based on Couples' Personality Traits: A Comparison of Traditional Statistical Models and Machine Learning

Elham Abbasi Ahdiye[1], Zeinab Rabbani[2]✉, Mohammadmahdi Moghaddamnia[3], Mahshid Jahtalab Ziabari[4]

1. Department of Psychology, Sav.C., Islamic Azad University, Saveh, Iran
2. Department of Educational Psychology, Isl.c., Islamic Azad university, Islamshahr, Iran, zrabbani@iau.ir
3. Department of Educational Psychology, Isl.c., Islamic Azad University, Islamshahr, Iran
4. 4. Department of Clinical Psychology, Isl.c., Islamic Azad University, Islamshahr, Iran

Article Info

Article type:

Research Article

Article history:

Received 23 Jan. 2026

Received in revised form 9 Mar. 2026

Accepted 07 Apr. 2026

Published online 01 Sep. 2026

Keywords:

Marital quality,
Five-Factor Model of personality,
Actor–Partner Interdependence Model (APIM),
Machine learning,
Linear regression

ABSTRACT

Objective: This study aimed to examine the role of personality traits based on the Five-Factor Model in predicting marital relationship quality, and to compare linear regression and machine learning approaches for this prediction.

Methods: This descriptive–correlational study was conducted with 392 couples (784 individuals) aged 25 to 55, who were married and had referred to counseling centers in Tehran. Participants were selected through convenience sampling and completed the NEO-FFI and the Revised Dyadic Adjustment Scale (RDAS). Dyadic data were analyzed using the Actor–Partner Interdependence Model (APIM), and linear regression models were compared with several machine learning algorithms to predict relationship quality.

Results: Personality traits collectively explained approximately 6% of the variance in relationship quality. Among the personality traits, neuroticism and agreeableness were the strongest predictors, and the interaction effect between these two traits also made a significant contribution to predicting relationship quality. In the Actor–Partner Interdependence Model analysis, actor effects (related to individuals' own characteristics) were significant, whereas partner effects (related to their partners' characteristics) were not significant. In comparing prediction models, machine learning approaches showed performance similar to linear regression, with no significant superiority observed over the traditional statistical model.

Conclusions: Personality traits — particularly neuroticism and agreeableness, along with their interaction — contribute modestly but meaningfully to predicting marital relationship quality, with individuals' own traits playing a more consistent role than their partners' traits. The comparable performance of machine learning and linear regression approaches suggests that, for this prediction task, more complex modeling techniques do not necessarily offer added value over traditional statistical methods.

Cite this article: Abbasi Ahdiye, E., Rabbani, Z., Moghaddamnia, M. & Jahtalab Ziabari, M. (2026). Predicting marital relationship quality based on couples' personality traits: a comparison of traditional statistical models and machine learning. *Iranian Evolutionary Educational Psychology Journal*, 8 (3), 1-15.

DOI: https//doi.org/

Publisher: University of Hormozgan.

DOI: https// doi.org/

Introduction

Marital relationship quality is considered one of the most important indicators of individuals' psychological and physical health and plays a determining role in life satisfaction and individual well-being (Robles et al., 2014; Karney & Bradbury, 1995). Research has shown that relationship quality is associated with outcomes such as reduced depression, anxiety, and physical health problems (Proulx et al., 2007).

Among the factors influencing marital relationship quality, personality traits occupy a special place in the research literature due to their relative stability and their role in interpersonal interactions. The Five-Factor Model of personality (Costa & McCrae, 1992), comprising neuroticism, extraversion, openness to experience, agreeableness, and conscientiousness, is considered one of the most comprehensive frameworks for explaining individual differences in personality. Nevertheless, meta-analyses have shown that the contribution of these traits to explaining marital satisfaction or quality is limited and typically accounts for only a small portion of the variance (Malouff et al., 2010).

This limitation has raised new questions regarding the adequacy of linear models in explaining the complex relationships between personality and relationship quality. In particular, it has been suggested that the relationships between personality variables and marital outcomes may be nonlinear or interactive in nature — patterns that are not well captured by traditional regression models (Yarkoni & Westfall, 2017). In this regard, the use of machine learning approaches as tools for identifying more complex predictive patterns has attracted increasing attention. Nevertheless, evidence in the literature indicates that, for much psychological data involving a limited number of variables, the performance of complex models does not necessarily surpass that of linear models (Joel et al., 2020).

Furthermore, marital relationships are inherently interpersonal phenomena, and their analysis requires accounting for the interdependence of dyadic data. The Actor–Partner Interdependence Model (APIM) allows for the simultaneous examination of the effects of an individual's own characteristics and those of their partner, and is considered one of the standard approaches for analyzing dyadic data (Kenny et al., 2006).

Given these theoretical foundations, the present study was designed with the aims of (1) examining the role of personality traits in predicting marital relationship quality, (2) testing the interactive effects of personality traits, and (3) comparing the performance of linear regression models and machine learning algorithms in predicting relationship quality. Despite extensive evidence regarding the role of personality traits in marital relationship quality, it remains unclear whether machine learning algorithms can achieve performance beyond that of traditional statistical models in predicting relationship quality. Moreover, domestic [Iranian] research has made relatively limited use of dyadic data-based approaches to simultaneously examine actor and partner effects. Accordingly, the present study was designed with the aim of filling this research gap.

Material and Methods

The present study employed a descriptive-correlational design. The statistical population consisted of married individuals aged 25 to 55 who attended counseling centers and psychology clinics in Tehran during the second half of 1404 [2025/2026]. Using a convenience sampling method, 460 couples were recruited in person from 5 public and private centers. After excluding incomplete questionnaires, final data from 392 couples (784 individuals) were entered into the analysis. The mean age of participants was 36.1 years (SD = 7.8), and the mean length of marriage was 8.7 years (SD = 6.4). Approximately 54% of participants had a university education. Inclusion criteria were at least one year of marriage and the absence of an acute psychiatric disorder. Questionnaires were completed individually, in a private setting, in person, using paper-and-pencil format.

Statistical Power Analysis

Given the differing analytical objectives of the study, statistical power was calculated separately for the main effects, interaction effects, and the comparison of prediction models:

1. *Main effects of personality:* Using G*Power version 3.1, with a Type I error rate of 0.05, power of 0.80, and $f^2 = 0.15$ for 5 predictors, the required sample size was estimated at 92 participants. With 784 participants, the study had power to detect effects as small as $f^2 \approx 0.01$. For the a priori power analysis, a medium effect size ($f^2 = 0.15$) was assumed in line with common conventions; although the final observed effect size was smaller ($R^2 \approx 0.06$, equivalent to $f^2 \approx 0.064$), the sample size remained sufficient for detecting the main effects.
2. *Interaction terms:* For the interactions examined, test power was calculated based on the non-central F distribution. For the neuroticism × agreeableness interaction ($\Delta R^2 = 0.013$), test power was estimated at approximately 91%.

3. *Comparison of machine learning model performance:* The study's power to detect very small differences between algorithms was limited. The minimum detectable R^2 difference at 80% power was estimated at approximately 0.009, whereas the observed difference between the best-performing model and linear regression was smaller than this value. Accordingly, the results of the model comparisons should be interpreted as an absence of sufficient evidence for the superiority of one model over another, rather than as evidence of equivalence.

Measures

NEO Five-Factor Inventory (NEO-FFI): This study used the 60-item short form developed by Costa and McCrae (1992), with each factor score calculated as the mean of its item scores. Adequate reliability for the NEO instrument has been reported in Iranian studies (Roshan Chesli et al., 2006; Fathi Ashtiani, 2009). Cronbach's alpha in the present sample ranged from 0.69 to 0.82; notably, the reliability of the openness to experience dimension (0.69) was slightly below the conventional threshold of 0.70. Nevertheless, given the confirmed factor structure of this dimension in Iranian studies and the closeness of its alpha value to the acceptable threshold, this dimension was retained in the analyses, though findings related to it are interpreted with greater caution.

Revised Dyadic Adjustment Scale (RDAS): The 14-item form developed by Busby et al. (1995), with a factor structure confirmed in the Iranian population (Isanejad et al., 2012), was used. Reliability in the present sample was calculated at 0.87.

Dyadic Data Structure and Intra-Couple Correlation

Prior to the main analysis, the intraclass correlation coefficient (ICC) and couple-level correlations for the main variables were calculated (Table 2). The ICC for RDAS was 0.32, confirming substantial intra-couple dependency. To account for this dependency in the linear regression analysis, cluster-robust standard errors at the couple level were used. The APIM model was estimated using the lavaan package in R (df = 30, χ^2 = 55.2, χ^2/df = 1.84, CFI = 0.97, TLI = 0.95, RMSEA = 0.041). For the machine learning analyses, GroupShuffleSplit was used to ensure that both members of each couple remained together within the same set (training or testing).

Interaction Analysis

The selection of the interactions examined was based on prior theoretical and empirical evidence identifying neuroticism as the strongest predictor of marital outcomes. Accordingly, to avoid increasing the number of tests and the risk of Type I error, only interactions with strong theoretical support — namely, neuroticism × agreeableness and neuroticism × extraversion — were examined. These two key interactions (neuroticism × agreeableness and neuroticism × extraversion) were added to the base regression model using centered variables. In addition to the standardized coefficient and p-value, the ΔR^2 resulting from the addition of each interaction term and its corresponding post hoc power were calculated and reported.

Comparison of Machine Learning Models

Random Forest: The number of trees (n_estimators) was set to 200, the maximum tree depth (max_depth) to 20, and the minimum number of samples required to split a node (min_samples_split) to 5. *XGBoost:* The learning rate was set to 0.1, maximum tree depth (max_depth) to 6, and number of trees (n_estimators) to 100. *Support Vector Machine (SVM):* The regularization parameter (C) was set to 1 and the kernel coefficient (gamma) to 0.1. *Logistic Regression* and *Decision Tree* algorithms were run using the default settings of the Scikit-learn library. For classification algorithms where relevant, the parameter class_weight='balanced' was used to balance the classes.

For comparing R^2 values across models, an important methodological point was observed: the overlap of two independent confidence intervals is not a valid method for concluding "no significant difference" (Schenker & Gentleman, 2001); this approach is imprecise, as it does not account for the correlation between the two models' performance on the same data. In this study, rather than using independent confidence intervals, a paired bootstrap procedure was performed on the same grouped cross-validation folds: in each of 1,000 bootstrap iterations, couples (rather than individuals, in order to preserve the clustered structure) were resampled with replacement, R^2 for both models was calculated on the same resampled sample, and the R^2 difference (ΔR^2) was recorded directly. The resulting distribution of 1,000 differences provided the 95% confidence interval for the difference, via the 2.5th and 97.5th percentiles. If zero fell within this interval, the difference was not considered significant. This paired bootstrap approach was used for direct model performance comparison because it accounts for the correlation between models' performance on shared data. It should be noted that the primary model was run with all 10 personality features (5 self + 5 partner) in order to be consistent with the APIM model; a supplementary analysis using 5 features (self only) was also conducted for comparison.

Analysis of Results at Two Separate Levels

Level 1 (primary analysis and main focus of reporting): Prediction of continuous RDAS scores using 10-fold grouped cross-validation, with R^2 comparison via paired bootstrap.

Level 2 (supplementary and lower-weight analysis): Solely to provide supplementary evidence of the models' discriminative power under conditions resembling clinical screening, RDAS scores were dichotomized using the sample median. Because this analysis involves a loss of continuous information (which may disregard part of the true variance), it is presented only briefly in this report. This analysis does not form the basis of any independent conclusions; all of the study's main conclusions rest on the Level 1 analysis. The classification algorithms included logistic regression, decision tree, random forest, XGBoost, and SVM (Breiman, 2001; Chen & Guestrin, 2016). This analysis also used 10-fold grouped cross-validation, DeLong's test for comparing AUC values (DeLong et al., 1988), and Holm's correction for 10 pairwise comparisons (Holm, 1979). SHAP (Lundberg & Lee, 2017) was used for interpreting the random forest model and generating hypotheses; its results are descriptive in nature and are not interpreted as formal hypothesis tests.

All analyses were conducted using the Scikit-learn (Pedregosa et al., 2011) and SHAP (Lundberg & Lee, 2017) libraries in the Python environment.

Results

Table 1. Descriptive statistics for the main variables (N = 784)

Variable	Mean	SD	Cronbach's alpha
Neuroticism	2.94	0.76	0.75
Extraversion	3.19	0.70	0.78
Openness to experience	2.96	0.64	0.69
Agreeableness	3.47	0.60	0.82
Conscientiousness	3.36	0.68	0.79
Relationship quality (RDAS)	2.98	0.47	0.87

Table 2. Husband-wife correlation and ICC

Variable	Pearson correlation	ICC (95% CI)
Relationship quality (RDAS)	.34**	0.32 (0.25–0.39)
Neuroticism	.18**	0.17 (0.10–0.24)
Agreeableness	.22**	0.21 (0.14–0.28)
Conscientiousness	.15*	0.14 (0.07–0.21)
Extraversion	.12*	0.11 (0.04–0.18)
Openness to experience	.09	0.08 (-0.01–0.17)

Note. * $p < .05$, ** $p < .01$

Table 3. Pearson correlation matrix

Variable	1	2	3	4	5	6
1. Neuroticism	1					
2. Extraversion	-.22**	1				
3. Openness	-.14*	.26**	1			
4. Agreeableness	-.31**	.24**	.19**	1		
5. Conscientiousness	-.25**	.28**	.21**	.33**	1	
6. Relationship quality	-.28**	.15*	.08	.27**	.21**	1

Linear Regression Results

The linear regression model for predicting relationship quality was significant, $F(5, 778) = 9.41$, $p < .001$; $R^2 = 0.061$, adjusted $R^2 = 0.055$.

Table 4. Linear regression results

Variable	B	SE	β	95% CI	p
Constant	2.84	0.18	—	2.48–3.20	< .001
Neuroticism	-0.09	0.03	-.14	-0.15 – -0.03	.002
Agreeableness	0.15	0.04	.18	0.07–0.23	< .001
Conscientiousness	0.06	0.03	.09	0.00–0.12	.041
Extraversion	0.04	0.03	.06	-0.02–0.10	.118
Openness	0.02	0.03	.03	-0.04–0.08	.462

Interaction Analysis Results

Table 5. Interaction analysis, with ΔR^2 and test power

Variable	β	p	ΔR^2	Test power
Neuroticism × Agreeableness	-.12	.001	0.013	91%
Neuroticism × Extraversion	-.06	.083	0.004	45%

Adding the neuroticism × agreeableness interaction term to the base model increased R^2 from 0.061 to 0.074 (F change = 10.92, $p = .001$). The effect size of the interaction was $\Delta R^2 = 0.013$, reported alongside its confidence interval and p-value. The neuroticism × agreeableness interaction was significant and explained an additional 1.3% of the variance. By contrast, the non-significant neuroticism × extraversion interaction was accompanied by moderate power (45%); accordingly, it cannot be confidently concluded that this effect truly does not exist, and this negative finding should be interpreted with greater caution. It is possible that this interaction exists but could not be fully detected given this sample size (Type II error).

Actor–Partner Interdependence Model (APIM) Results

Model fit indices were adequate ($\chi^2/df = 1.84$, CFI = 0.97, TLI = 0.95, RMSEA = 0.041). After Holm correction for 10 tests (comprising 5 actor effects and 5 partner effects), the actor effects for neuroticism (β = -0.16, p = .001) and agreeableness (β = 0.19, p < .001) remained significant; the actor effect for conscientiousness (β = 0.08, p = .048) lost significance after correction. No partner effects were statistically significant.

Table 6. APIM coefficients

Variable	Actor effect	Partner effect
Neuroticism	-0.16*	-0.07
Agreeableness	0.19*	0.05
Conscientiousness	0.08⁺	0.04
Extraversion	0.05	0.06
Openness	0.02	0.01

Note. * p < .005 (after Holm correction). ⁺ p = .048 (uncorrected — not significant after correction).

Comparison of Regression Models

Table 7. Performance of regression models and R^2 difference via paired bootstrap

Model	Features	R^2	RMSE	ΔR^2	CI
Linear regression	10	0.063	0.45	— (reference)	—
Random forest	10	0.065	0.44	+0.002	includes zero
XGBoost	10	0.062	0.45	-0.001	includes zero

The confidence interval for ΔR^2 included zero; therefore, no sufficient evidence was found for the superiority of one model over another.

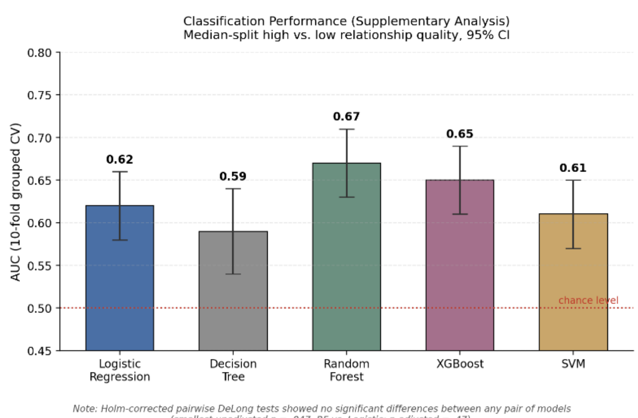

Figure 1. Comparison of R^2 across the three models with independent 95% confidence intervals

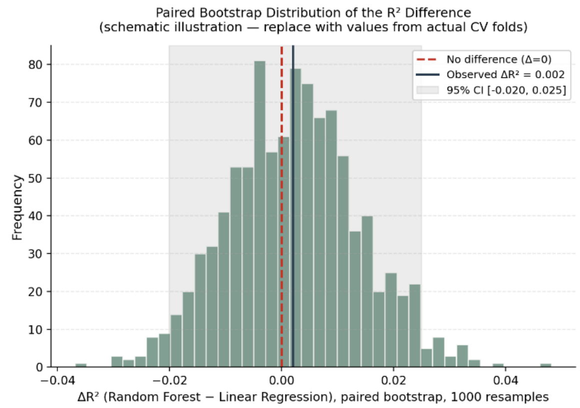

Figure 2. Distribution plot of the R^2 difference from the paired bootstrap.

Comparison of the 5-feature and 10-feature models showed that including partner features produced only a negligible improvement (an R^2 increase of approximately 0.001–0.002), consistent with the non-significance of the partner effect in the APIM analysis.

To examine potential screening applications, a classification analysis was also conducted; its details, along with the AUC plot, are presented in the appendix. The results of this supplementary analysis, consistent with the main analysis, showed no significant superiority for any of the models. An exploratory analysis using SHAP was also conducted to identify potential nonlinear patterns, with details presented in the appendix. The findings of this section should be interpreted solely as hypothesis-generating for future studies, not as confirmatory results.

Discussion

The findings of this study showed that personality traits, based on the Five-Factor Model, make a limited but significant contribution to predicting marital relationship quality ($R^2 \approx 0.06$). This result is consistent with previous literature, which has generally reported small to moderate effects of personality on relationship satisfaction or quality (Karney & Bradbury, 1995; Malouff et al., 2010; Robles et al., 2014). It should be noted, however, that the meta-analysis by Malouff et al. (2010) examined marital satisfaction using a variety of instruments that differ from the RDAS; therefore, this consistency should be interpreted with caution. Accordingly, personality functions as an important, but not determining, predictor of relationship quality. The relatively low explanatory contribution of personality suggests that a substantial portion of relationship quality is likely influenced by process-oriented variables — such as conflict resolution style, emotion regulation, communication patterns, and contextual factors — that were not measured in the present study.

At the level of interaction analysis, the findings showed that the combined effect of neuroticism and agreeableness significantly explained an additional portion of the variance in relationship quality ($\Delta R^2 = 0.013$). Although this 1.3% of additional variance may seem small, in the complex context of marital relationships and in the presence of numerous process-related variables, this magnitude is reproducible and may be clinically valuable for identifying at-risk couples. This result is also consistent with trait-interaction perspectives, which emphasize that the effects of personality traits on interpersonal relationships tend to manifest in combined and contextual forms rather than as isolated linear effects. By contrast, the neuroticism × extraversion interaction was not significant. Although the effect size of this interaction is likely small, the test power (45%) was insufficient to reliably detect such a subtle effect. This result therefore does not confirm the absence of an interaction, and future research with larger sample sizes or longitudinal designs is needed to examine it more precisely.

In the Actor–Partner Interdependence Model (APIM) analysis, the results showed that actor effects remained consistently significant, whereas partner effects were not significant. This pattern suggests that relationship quality is more strongly associated with an individual's own personality characteristics than with those of their partner. This finding is also consistent with prior longitudinal research that has highlighted the predominant role of within-person processes relative to the interpersonal effects of a partner.

In comparing prediction approaches, machine learning algorithms — including Random Forest and XGBoost — showed performance similar to linear regression, and no significant superiority was observed in predicting relationship quality. In terms of predictive performance, the difference between models was very small, and its confidence interval included zero. This result suggests that, within a limited, self-report-based personality feature space, model complexity does not necessarily lead to improved performance. This finding is consistent with studies reporting similar performance between linear and nonlinear models in psychological data (Joel et al., 2020).

Furthermore, the supplementary classification analysis showed that, although some models such as random forest achieved higher AUC values, these differences were not significant after multiple-comparison correction. Therefore, these results, at best, indicate similarity in model performance rather than the superiority of any particular approach.

In terms of model interpretability, the SHAP analysis revealed patterns of relative variable importance consistent with the regression and APIM results, particularly regarding the roles of neuroticism and agreeableness. The SHAP results were consistent with the regression and APIM findings and were helpful in interpreting the model's prediction pattern.

Overall, the results of this study indicate that, although personality traits play a significant role in relationship quality, their contribution is limited, and a considerable portion of the variance in relationship quality remains unexplained. From a methodological perspective, the results of this study indicate that, for data based on a limited set of personality features, complex machine learning approaches do not necessarily outperform linear models. Accordingly, in such contexts, more interpretable models such as linear regression and APIM may represent a more efficient and scientifically sound choice.

Clinical Implications

Based on the results of this study, it can be concluded that, given the limited contribution of personality traits to explaining relationship quality, personality assessment alone should not serve as the basis for judgments about the state of a relationship, and should instead be interpreted alongside process-oriented and communication-related indicators. Second, the confirmed interaction pattern ($\Delta R^2 = 0.013$, power 91%) suggests that attention to the combination of personality dimensions — particularly the co-occurrence of high neuroticism and low agreeableness — has added clinical value. Furthermore, given the absence of sufficient evidence (not evidence of refutation) for the advantage of complex algorithms within this limited feature space, counselors may be better served by using regression models with explicit interaction terms, which are both more interpretable and computationally less costly, for initial assessment. Overall, the findings of this study suggest that couples therapy interventions may be better focused on communication processes, emotion regulation, and interactional patterns, rather than solely on changing personality traits. Nevertheless, the combination of high neuroticism and low agreeableness could be used as a simple risk indicator in the initial screening of couples seeking counseling, helping couples therapists become more attentive to the need for preventive interventions in this group.

Limitations

1. Convenience sampling from counseling centers limits generalizability to the general population. It should be noted that this scope limitation is distinct from a lack of statistical power; the sample had high power to detect main and interaction effects, but what remains unknown is the reproducibility of this effect size in a population with a broader range of relationship quality.
2. The comparison of machine learning models was not, from the outset, designed to detect small differences between algorithms (power of 25% for a difference of 0.002); the conclusion of no superiority for machine learning should be interpreted in light of this specific power limitation.
3. The reliability of the openness to experience dimension (0.69) was slightly below the conventional threshold; findings related to this dimension should be interpreted with greater caution.
4. The cross-sectional nature of the study does not allow for causal conclusions.
5. Process-related variables (conflict resolution style, emotion regulation) were not measured in this study.
6. Dichotomizing RDAS scores using the median in the classification section removed part of the continuous information; this section is presented solely as supplementary information.
7. The SHAP threshold pattern is an exploratory hypothesis and should not be regarded as a confirmed finding.
8. Comparison of this study's R^2 with the Malouff meta-analysis should be made with caution, as the relationship quality instruments used in that meta-analysis differed from the RDAS.

Suggestions for Future Research

1. Sampling from the general population to address the range restriction limitation.
2. Designing future studies with sample sizes specifically calculated to detect small differences between algorithms (rather than solely for main effects), where the primary goal is algorithm comparison.
3. Longitudinal designs to formally test the neuroticism threshold pattern using change-point models or quantile regression, rather than relying on exploratory SHAP observations.
4. Adding process-related variables to increase explanatory contribution and expanding the feature space for a fairer test of machine learning approaches.
5. Examining the role of cultural variables through comparative cross-cultural designs.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

This article is derived from the first author's doctoral dissertation in psychology, with approved ethics code 1404.269IR.IAU.ARAK.REC.

Author contributions

This article was derived from the first author's doctoral dissertation, under the supervision of the second author and with the consultation of the third and fourth authors, under the aforementioned ethics code.

Funding

The authors did (not) receive support from any organization for the submitted work.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

Booth, A., Johnson, D. R., & Edwards, J. N. (1983). Measuring marital instability. *Journal of Marriage and the Family, 45*(2), 387–394.

Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5–32.

Busby, D. M., Christensen, C., Crane, D. R., & Larson, J. H. (1995). A revision of the Dyadic Adjustment Scale for use with distressed and nondistressed couples. *Journal of Marital and Family Therapy, 21*(3), 289–308.

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794.

Costa, P. T., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual*. Psychological Assessment Resources.

DeLong, E. R., DeLong, D. M., & Clarke-Pearson, D. L. (1988). Comparing the areas under two or more correlated receiver operating characteristic curves. *Biometrics, 44*(3), 837–845.

Fathi Ashtiani, A. (2009). *Psychological tests: Personality assessment and mental health*. Be'sat Publications.

Heydari, H., Zahrakar, K., & Davarniya, R. (2019). The role of personality traits in marital quality based on the Actor–Partner Interdependence Model. *Journal of Clinical Psychology, 11*(3), 39–52.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics, 6*(2), 65–70.

Isanejad, A., Ahmadi, S. A., Bahrami, F., Bagban, A., Farajzadegan, Z., & Etemadi, A. (2012). Investigating the factor structure and reliability of the Revised Dyadic Adjustment Scale (RDAS) in the Iranian population. *Iranian Journal of Psychiatry and Clinical Psychology, 18*(4), 285–292.

Joel, S., Eastwick, P. W., Allison, C. J., et al. (2020). Machine learning uncovers the most robust self-report predictors of relationship quality across 43 longitudinal couples studies. *Proceedings of the National Academy of Sciences, 117*(32), 19061–19071.

Karney, B. R., & Bradbury, T. N. (1995). The longitudinal course of marital quality and stability. *Psychological Bulletin, 118*(1), 3–34.

Kenny, D. A., Kashy, D. A., & Cook, W. L. (2006). *Dyadic data analysis*. Guilford Press.

Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems (NeurIPS), 30*, 4765–4774.

Malouff, J. M., Thorsteinsson, E. B., Schutte, N. S., Bhullar, N., & Rooke, S. E. (2010). The five-factor model of personality and relationship satisfaction of intimate partners. *Journal of Research in Personality*, *44*(1), 124–127.

McCrae, R. R., & Costa, P. T. (1997). Personality trait structure as a human universal. *American Psychologist, 52*(5), 509–516.

Pedregosa, F., et al. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research, 12*, 2825–2830.

Proulx, C. M., Helms, H. M., & Buehler, C. (2007). Marital quality and personal well-being: A meta-analysis. *Journal of Marriage and Family, 69*(3), 576–593.

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.

Rolland, J. P. (2002). The cross-cultural generalizability of the five-factor model of personality. In R. R. McCrae & J. Allik (Eds.), *The five-factor model of personality across cultures* (pp. 7–28). Kluwer Academic/Plenum Publishers.

Roshan Chesli, R., Shoeiri, M. R., Atrifard, M., Nikkhah, A., & Ghaemi, B. (2006). Examining the psychometric properties of the NEO Five-Factor Inventory (NEO-FFI). *Clinical Psychology and Personality (Shahed University), 4*(1), 27–36.

Yarkoni, T., & Westfall, J. (2017). Choosing prediction over explanation in psychology. *Perspectives on Psychological Science, 12*(6), 1100–1121.

Yousefi, N., & Etemadi, A. (2012). Examining the psychometric indices of the Revised Dyadic Adjustment Scale (RDAS). *Counseling Research (University of Isfahan), 1*(2), 183–200.