Who Makes Whom Happy in a Romantic Relationship?
Contributions to the Predictive Validity of Personality Questionnaires
in Mating Contexts.

Doctoral Dissertation

University of Hamburg

Faculty of Psychology and Human Movement

Department of Psychology

Von-Melle-Park 5, 20146 Hamburg, Germany

Inga Grossmann

Matriculation Number: 6483333

Email: inga.grossmann@uni-hamburg.de

Date of Submission: 25.09.2017

Date of Disputation: 19.03.2018

Vorsitzender des Prüfungsausschusses:	Prof. Dr. Alexander Redlich
Erster Dissertationsgutachter:	Prof. Dr. Burghard Andresen
Zweiter Dissertationsgutachter:	Prof. Dr. Jan Wacker
Erste Disputationsgutachterin:	Prof. Dr. Monika Luzi Beyer
Zweite Disputationsgutachterin:	Prof. Dr. Juliane Degner

Abstract

Whose romantic relationships last happily? The present work aimed to contribute to the predictive validity of personality questionnaires in mating contexts. Thereby, it refined the understanding of how personality and romantic relationships - particularly their quality - interplay. Five studies that utilized questionnaires for self-assessing both relationship-related and *general personality traits* were conducted.

Study I and *IV* focused on predicting *relationship quality* with the help of methods from *machine learning*: in both, linear additive models were developed and cross-validated based on different sets of personality variables as predictors. They reproducibly predicted relationship quality to an unprecedented extent. In the four-year longitudinal *Study I*, variables of *relationship-related personality* better than those of *general personality* predicted future *relationship quality*, while actor effects were better predictors than partner, similarity, and other interaction effects. In contrast to *Study I*'s findings, interaction effects (fittings, similarities, other moderators) outperformed actor and partner effects in the cross-sectional *Study IV*; this may be due to the usage of a novel matching test which was developed and validated successfully in *Study III,* thus paving the way to be applied in *Study IV* and *V*. The novel economic questionnaire measures one's own and preferred partner characteristics in a parallel design to enable real-ideal fittings with a current (potential) partner. Just as similarity scores, fitting scores showed high predictive validity in *Study III,* but were not able to incrementally contribute to prediction of *relationship quality* in the cross-sectional *Study IV*.

Requesting transferability to singles in mating contexts, *Study II* and *V* contributed to assess the cross-contextual applicableness of prediction models, which are typically developed on couple's datasets in anonymous settings. During the four-year-longitudinal *Study II*, relevant changes of *relationship-related personality* facets over time and different relationship statuses were tested and described. Unsurprisingly, but still pioneering, many facet changes were relevantly affected by current partnerships and their outcome, while others were not. The findings partly limit the applicableness of models from *Study I*, which are based on the same personality test. In *Study V*, answer distortions between an anonymous and a fake online dating setting were tested and described. Patterns of self-exaggeration, lower retest-correlations for preferred than for own characteristics, and setting-related differences in predictive validity regarding *relationship quality* were found. These outcomes partly limit the applicableness of models from *Study IV*, which are based on the same personality test.

Summarized Table of Content

Complete Table of Content

1. *General Introduction*

1.1. *Background*

1.1.1. *Relevance of the topic*

For most adults, a central goal in life is to attain and to maintain a satisfying romantic relationship; this plays a key role in fostering their well-being (Berscheid, 1999). A review by Kiecolt-Glaser and Newton (2001) and a meta-analysis by Proulx, Helms and Buehler (2007) showed moderate cross-sectional and longitudinal correlations of RQ (*relationship quality*) to mental and physical health. However, why are some relationships successful and satisfying, while others have a negative impact on health? A study by Solomon and Jackson (2014), using a representative and longitudinal sample, suggested that the personality of partners influences *relationship satisfaction*, which in turn influences the likelihood of a *break-up*. As most *personality traits* are stable across different relationships, this naturally leads to the question, if they can be used to predict the RQ of a possible future couple. If so, this could allow for forms of matchmaking that increase RQ and, therefore, the well-being of both partners.

Through omnipresent internet access, increasing social acceptance, and general efficiency in material and time costs, online dating has broadly evolved to become a feasible and common practice in the search for potential long-term partners, particularly in Western-orientated areas (DrexelUniversity, 2012). Online matching algorithms predominantly rely on the similarity of partner's personality (Bartholomew & Horowitz, 1991; Hitsch, Hortaçsu, & Ariely, 2010). Many portals, such as Parship, ElitePartner or OkCupid, claim to use effective, consolidated dyadic matching algorithms. However, even in an extensive search, no online dating website which employs cross-validated models that accurately predict RQ longitudinally was found. Indeed, a recent study (Cacioppo, Cacioppo, Gonzaga, Ogburn, & VanderWeele, 2013) using a nationally representative American sample of partners, who were married between 2005 and 2012, indicated that more than 35% of marriages started online; 45% of these partners stated they met through online dating websites, indicating its high feasibility as a modern mating context. However, online dating has not led to better partner matching than offline dating thus far: marriages that started online, only exceedingly slightly less resulted in *break-up* and lower *relationship satisfaction* than marriages which started offline. This indicated that the current algorithms which try allocating potential part-

ners to one another, who would more probably make one another happy in a future romantic relationship, do not work properly. The reasons why the partner matching algorithms do not work has not yet been elaborated. Thereby, the question of how and to which extent partner matching can enhance future couple's RQ in such real-life mating contexts is still a mainly unanswered research question.

1.1.2. Reproducible success of previous prediction models

Existing research has already addressed the question of to what extent it is possible to predict RQ based on personality. However, previous approaches working with similarity, actor and partner variables mostly used simple correlational approaches, such as structural equation-based modelling, and generally found only modest effects (Finkel, Eastwick, Karney, Reis, & Sprecher, 2012). These regular explanatory approaches may fail when facing the complexity of a real interplay between the different traits of two individuals.

Some approaches using mathematically more sophisticated models optimized predictive replicative power for *break-up* (Gottman & Notarius, Decade review: Observing marital interaction, 2000) based on characteristics of marital interaction in a present partnership such as communication, conflict, and mood aspects (Gottman, 2014). For instance, an accurate model was developed using a discriminant analysis in a 10-fold CV (*cross-validation*) based the test system ENRICH. It predicts a *break-up* with a longitudinal accuracy of 80 – 90% (Fowers & Olsen, 2007), but only functions properly for existing relationships. Methods based exclusively on the highly stable *personality traits* of the partners could, in contrast, also be used to predict the RQ of a potential future couple. However, until now, the question is left open if personality traits not only reproducibly predict initial romantic attraction (Joel, Eastwick, & Finkel, 2017) but also later RQ. The present work addressed this question with *Study I* and *IV*. By using methods of ML (*machine learning*), these followed the recent interdisciplinary methodological trend in the field of cognitive and social psychology (Rosenfeld, Zuckerman, Azaria, & Kraus, 2012; Yarkoni, Ashar, & Wager, 2015; Youyou, Kosinski, & Stillwell, 2015).

Yarkoni and Westfall (2016) have discussed that short-term emphasis on reproducible prediction could ultimately improve the ability to explain the causes of behavior over the long term and, thereby, increase theoretical understanding. Their recent work has shown that ML methods can contribute to solving the problem concerning the reproducibility of a researcher's analysis: Traditional methods of analyzing data in the field of psychology follow an explanatory pattern and lead to issues such as overfitting to specific data sets by the evaluation

procedure (Lucas & Donnellan, 2013; Open Science Collaboration, 2015). So-called 'P-hacking' (Simmons, Nelson, & Simonsohn, 2011) or less tendentiously, data-contingent analysis (Gelman & Loken, 2013) is one of the most common causes of overfitting biases in psychological research; it is particularly relevant for small, non-representative data sets.

ML applies advanced methods to train prediction models which can create reproducible results. The avoidance of overfitting during the model training is a primary goal (Domingos, 2012). Specialized variable engineering and variable selection can cope with special characteristics and restrictions of a given data set, such as many highly correlated variables and small sample sizes, as it is often the case in psychological studies (Yarkoni & Westfall, 2016). In ML, the success of a research result is defined as so: 'success' is not measured by the degree of a model's fit (e.g., a theoretically privileged regression coefficient) for the data the model is developed on, but instead by the average difference between unobserved data and the model's predictions for those new data. To conclude, model fitting and model evaluation should, under no circumstances, be carried out on the same dataset. To comply with this claim, *Study I* and *IV* have employed cross-validating designs.

1.1.3. Accuracy- and application-related personality test evaluation

For matching purposes, the characteristics of both partners were mostly measured by questionnaires for self-assessment. These diagnostics typically are accompanied by certain accuracy-related complications as so regarding the following three topics:

- **Stability**

  Self-assessed characteristics which adapt to different relationship statuses and relationships (Kirkpatrick & Hazan, 1994; Specht, Schmuckle, & Egloff, 2011), represent a considerable problem for matching tests. Matching calculations are mainly based on data from subjects who already have a relationship and, thereby, may partly not be valid for a new partnership. Moreover, personality aspects that change too much over time may be less suitable for long-term predictions, overall. The relationship-related long-term changeability of *relationship-related personality traits* has not yet been examined beyond attachment styles. Therefore, S*tudy II* addressed this topic.

- **Social desirable responding (SDR)**

  Previous research has shown the effects of impression management and its possible influences on practical psychological decision-making settings (Berry, Ones, & Sackett, 2007; Dilchert, Ones, Viswesvaran, & Deller, 2006). Mating contexts are

such *high-stake* situations, yet - so far - matching tests have not been quantitatively tested for effects of answer distortions and their possible influence on predictive validity for RQ. *Study V* has focused upon that topic.

- *Scaling*

  A matching test should measure *relationship-related personality*, as well as *general personality*, since both have relevant associations with RQ (Noftle & Shaver, 2006). Moreover, predictive accuracy may be enhanced by using differentiated and homogeneous facets, instead of larger and abstract domain scales. Finally, previous research has found compatibilities between ideal and real partner to be promising for RQ (Fletcher, Simpson, & Thomas, 2000; Gerdvilyte & Abhyankar, 2010; Zentner, 2005). Consequently, a matching test may profit from enabling to score such fittings. Based on these demands, a novel matching test was designed and evaluated in *Study III*.

1.2.   Description of the present work

1.2.1. Objectives and conducted studies

The present work addresses a core question: whose relationship with whom will last happily; hence, which partners will maintain satisfied relationships? Five studies were designed to answer this question by evaluating how accurately self-assessed personality predicts RQ and what influences these predictions. *Figure 1* offers an overview concerning the examined phenomenon, topics, and their relations within the current work.
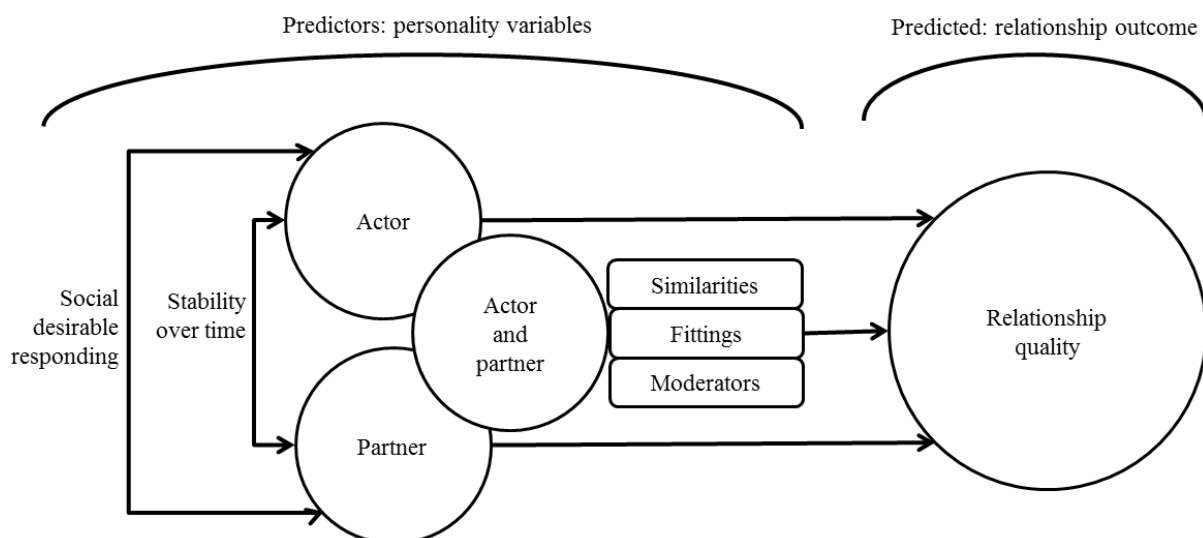
Figure 1. Model to predict *relationship quality* by *personality traits* of both partners.

Table 1 presents an overview of the studies' focus, designs, and samples. *Study I* and *IV* are directly focused on predicting RQ based on personality models with ML. *Study II*, *III* and *V* addressed relevant accuracy-related topics regarding the transferability of such models to real-life mating contexts.

In *Study I*, ML and partner matching were paired for the first time. Longitudinal linear additive models were developed and cross-validated, based on different sets of variables, including actor, partner, similarity and hypotheses-based moderator variables, as well as aspects from different personality domains. The models predicted different *RQ measures* such as *Sexual satisfaction, relationship satisfaction*, *separation (intents)*, and *conflicts*. Additionally, the predictive validity of a novel scoring which scales different kind of the partner's (dis)similarities was evaluated.

In *Study II*, the same facet scales of *relationship-related personality* (as in *Study I*), were tested for their stability over time and over different relationship surroundings (absent, ending, ongoing and new partnerships). Patterns of facet changes were described, as was partnership-overarching, and more adaptive *personality traits* were determined.

In *Study III,* a novel-matching test was designed and evaluated. Test economy, applicability in professional real-life mating contexts, and substantial contents, which are potentially related to predictive validity for RQ, were considered. Thereby, the novel scoring was applied and evaluated for (non)fittings, as well as for (dis)similarities.

In *Study IV*, the test designed in *Study III* and the methods from *Study I* were used to develop and cross-validate models predicting RQ based on different sets of variables in a cross-section. Findings were conflated with those from *Study I*. Influences and interactions of culture, partner ideals, and ideal-real fittings on RQ were additionally evaluated.

In *Study V,* the novel test applied in S*tudy III* and *IV* was tested for SDR by comparing answers from an online dating context to those of an anonymous setting. Patterns of general and individual *answer biases,* personality correlates, and influences on predictive validity were examined, as well as an attempt for individual faker detection was tested.

The first two studies employed a four-year longitudinal design and worked with the extensive *Relationship- and Attachment-related Personality Inventory* - the so-called BB-PI (Andresen, 2012) - to measure *relationship-related personality*. In *Study I,* the *Personality Domain Inventory* - the so-called PD-I (Andresen, 2015) - was applied additionally to measure *general personality*. *Studies III, IV* and *V* worked with and on a more economic, novel-matching test, the *YOUME-L(ist)* (Andresen, 2012) - which measures both relationship-

related and *general personality*. All applied diagnostic tests are described in detail (see *2.2.1 Operationalization and 4.2.1 Operationalization).*

Table 1: Overview of the conducted studies in the present work

	Test validation	Focus	Design	Operationalization	N
Study I	Additive predictive validity	RQ prediction models with *machine learning*	Longitudinal over 4 years	Time 1: BB-PI, PD-I, Time 2: RQ	(476)192
Study II	Retest-reliability	Changes over time and different relationship surroundings	Longitudinal over 4 years	Time 1 and 2: BB-PI	283
Study III	Psychometric test evaluation	Scaling for preferred partner and own traits	Cross-sectional	2x74/40 YOUME-L RQ	423 (250)
Study IV	Additive predictive validity	RQ prediction models with *machine learning*	Cross-sectional	2x74 YOUME-L RQ	192
Study V	Retest-reliability	Social desirable responding in a dating setting	Cross-sectional within-subject	2x40 YOUME-L	309

Notes. BB-PI: Relationship- and Attachment-related Personality Inventory (Andresen, 2012). *PD-I: Personality Domain Inventory PD-I* (Andresen, 2015). *YOUME-L(ists): available as an earlier short version with 2x40 items* (Andresen, WPRP-L, 2012) *and as an extended version with 2x74 items. RQ: relationship quality.*

1.2.2. Data collection[1]

The present section describes how the recruiting of the samples for the conducted studies was interlinked, as well as which and why samples overlap. *Figure 2* shows an overview of the datasets which were used and are represented by the different boxes.

Box 1.: N=192
used in Study I
(predictions)
Data of both partners
T1: BB-PI, PD-I, RQ; T2: RQ

n=120 from Stern Survey
n=45 from Wunderlich Survey
n=27 whose partner attended the
Stern Survey (T1) without them

Overlap Box 1 and 2:
the n=120 who
the Stern Survey
with their partner

Box 2: N=283
used in Study II and III
(retest-stability)
Data of singles and partners
T1/T2: BB-PI (all n),
YOUME-L-Wish (n=250)
from Stern Survey

n=120 with their partners
n=46 without their partners
n=117 singles

Box 3: N=284
used in Study I
(predictions)
Data of both partners
T1 only: BB-PI, PD-I, RQ
from Stern Survey

Box 4: N=192
used in Study III and IV
(test evaluation and predictions)
Data of both partners
T1 only: private setting
YOUME-L, RQ
From Survey at University of
Hamburg

Overlap Box
4 and 5:
the n=78
of those who
attended with
their partner

Box 5: N=309
used in Study III and V
(test evaluation and answer distortions)
Data of singles and partners
T1 only: private and public setting (2x)
YOUME-L
From Survey at University of Hamburg

n=78 with their partner
n=84 singles
n=147 without their partner

Figure 2. Overview of the datasets used for the five studies and their overlaps.

Notes. Boxes represent the different datasets. The sizes of the boxes represent the sample sizes.
Study I: N=476 T1 (Box 1 + 3: 284 + 192); N=192 T1/T2 (Box 1: 120 + 45 + 27).
Study II: N=283 (Box 2: 120 + 46 + 117).
Study III: N=423 T1 (Boxes 4 + 5 – overlap: 192 + 309 - 78); N=250 T1/T2 (Box 2 – 33: 283 - 33).
Study IV: N=192 (Box 4).
Study V: N=309 (Box 5: 78 + 84 + 147).
Box 3 and T1 of Box 2 were recruited in the Stern Survey 2008.
Boxes 4/5 and T2 of Boxes 1/2 were recruited in surveys at the University of Hamburg.
Sample sizes of Boxes 1 and 4 are the same by chance: they were recruited in independent surveys and do not overlap.
Samples of Box 2 and 4/5 do not overlap: they are used for different analyses in Study III.
Sample sizes of Boxes 2 and 3 are nearly the same by chance and do not overlap: both were recruited in the Stern survey (T1) but samples from Box 2 participated in the T2 survey at the University of Hamburg while samples from Box 3 did not.
Only 250 of the 283 participants from Box 2 provided data for the Wish-Scales of the YOUME-L: therefore, only these were used for Study III which focused on the evaluation of the YOUME-L.
T1: Time 1. T2: Time 2, four years later. RQ: Relationship Quality.
BB-PI: Relationship- and Attachment-related Personality Inventory (Andresen, 2012).
PD-I: Personality Domain Inventory (Andresen, 2015).

---

[1] All collected data sets are stored on the appositive medium (Appendices I.1, II.1, III.1, IV.1 and V.1).

In 2008 (T1=Time 1), a sample of 1,218 subjects was recruited by a display corresponding with an article about partnerships in the popular German magazine *Stern*. The questionnaires included the entire item pool for the test construction of the BB-PI (Andresen, 2012), and the PD-I (Andresen, 2015) and optionally the short version of the YOUME-L (Andresen, WPRP-L, 2012); these were sent home to interested participants, who were required to send them back upon completion. Some respondents did not have a partner, some attended without, and some with their partner.

Four years later, in 2012 (T2=Time 2), approximately one-third of the 'contactable' sample from T1 was convinced to participate again in an online-survey conducted at the University of Hamburg by the author of this dissertation: 951 subjects had left a correct correspondence address at T1 and 324 of these could be recruited to participate in the T2 survey. On both occasions, the overall survey took between three to five hours, while small and long breaks were possible within the range of the duration of the investigation, which lasted for approximately three months: 41 subjects broke off the T2-survey prematurely. The sample was divided as follows:

- *Study I:* couples were exhaustively tested for their personality and RQ. The 284 partners, who only attended with their counterpart at T1 were included in the sample for cross-sectional analyses which mounted up to 476 partners (284 + 192). The 192 partners (120 + 27 + 45) who provided long-term data in some form were included in the longitudinal analysis. These samples were created as a patchwork of couple's data for individuals:
  - n=404: both partners' personalities at T1 were completed as part of the original *Stern* survey, n=120 of these provided RQ data at T2.
  - n=27: partners of those who participated in the *Stern* survey at T1 without them. These 27 partners only provided data at T2.
  - n=69: one or both partners did not take part in the *Stern* survey, but in a follow-up study one year later (Wunderlich, 2011), n=45 of them provided T2 RQ data.
- *Study II:* 283 participants of the T1 *Stern* survey completed the BB-PI again at T2. 120 of these attended with their partners, 117 were single and 46 attended without their partner at T1.
- *Study III*: 250 of these 283 subjects additionally completed the scales for partner preferences of the short YOUME-L at T1 and T2 - and, hence, could be used to examine longitudinal retest-reliability as subordinated part of the psychometric test validation.

Additional cross-sectional online surveys applying the YOUME-L were conducted at the University of Hamburg. In the realm of their thesis (Findeisen, 2014; Kaiser, 2014; Paki, 2014), three psychology students and the author of the current dissertation recruited participants among their fellows, relatives, and acquaintances. The following samples came about:

- *Study III:* 423 participants completed the long version of the YOUME-L and were used as the main sample for psychometric validation.

- *Study IV:* Among these 423, 192 individuals attended the study with their partners (96 couples): they were asked for their RQ and, thereby, built the basis for testing predictive validity among couples. 26 of these couples were Persian who completed a professionally translated Persian version of the YOUME-L[2] and were recruited by Setareh Paki.

- *Study V*: The 84 singles, the 147 subjects who attended the study without their partner, and the 78 subjects who already attended with their partner but were engaged to answer more questions, were instructed to fill out the short version of the YOUME-L twice: once anonymously and once in a faked online dating setting. Excluding dropouts, this data of 309 participants was used to examine SDR.

Full-online questionnaires with automatic data storing have the advantage that no errors can occur due to data input: e.g., by processes such as copying or merging data per hand. At the same time, the opportunity to store data and to continue later may have reduced the fatigue effect. However, the uncontrolled setting - i.e., participants being in full control of when, where, and how they answer online - could also have had negative consequences: e.g., on data consistency and standardizing answering conditions for the participants.

As an incentive for participation and honest responding, participants were given professional individual feedback on the results of their personality tests after each conducted survey. For studies at the University of Hamburg, individual assessments were programed and compiled in the *limesurvey* system, which automatically sent the results via email if requested[3].

---

[2] Appendix III.3. presents the Persian version of the YOUME-L and the survey instruction.
[3] Appendix III.4. presents an exemplary individual matching feedback.

1.2.3. Own and external contribution

Table 2 summarizes who took on which tasks and elements of the present work. The majority was conducted and taken care of by the author of this dissertation herself, as expected. The largest external support was: (1) The extensive reviews by Prof. Dr. Jan Wacker on early versions of *Study V's* and *II's* results, as well as (2) the support of André Hottung for the subjects of ML in *Studies I* and *IV*.

Table 2: Own and external contributions to the present work

	Study I	Study II	Study III	Study IV	Study V
Data collection	BA: Time 1 IG: Time 2	IG		IG, SP, KF, LK	
Data preparation	IG	IG	IG	IG	IG
Literature overview	IG	IG	IG	IG	IG
Design/ methods	IG: anything but AH: parameter tuning	IG: anything but JW: adaption	IG	IG	IG
Calculations	AH: code Python IG: code R, SPSS	IG	IG	AH: code Python IG: code R, SPSS	IG
Writing	IG: anything but AH: ML part of methods	IG	IG	IG	IG
Proofreading	AH JW	JW	JW		JW
	IG: final proofreading and correction with the help of https://www.deepl.com/translator				

Notes. *ML: machine learning. AH: André Hottung. JW: Jan Wacker. BA: Burghard Andresen. SP: Setareh Paki. LK: Anna Lucia Kaiser. KF: Kerstin Findeisen.*

2. **Study I***: Long-term Prediction of Relationship Quality with Machine Learning by Personality Traits*

Abstract

To what extent is it possible to use *machine learning* to predict the outcome of a relationship, based on the personality of both partners? In the present study, *relationship satisfaction*, *conflicts*, and separation (intents) of 192 partners four years following the completion of questionnaires concerning their *personality traits* was predicted. A 10x10-fold *cross-validation* was used to ensure that the results of the linear additive models could be reproducible. The findings indicate that *machine learning* techniques can improve the prediction of *relationship quality* (37% of variance explained), and that the perceived *relationship quality* of a partner is mostly dependent on his or her own individual *personality traits*.

Additionally, the influences of different sets of variables on predictions are shown: partner and similarity effects did not incrementally predict *relationship quality* beyond actor effects. *General personality traits* predicted *relationship quality* less strongly than *relationship-related personality traits* and the actor's past *tendencies to react emotionally positive or negative* to partnerships were the strongest predictor of future *relationship quality*.

Keywords: relationship satisfaction, relationship dissolution, *machine learning*, *personality traits*, dyadic data analysis, linear regression, *Big Five*, *cross-validation*, *personality traits*

2.1. Additional background[4]

2.1.1. Actor-, partner- and similarity effects

To which extent certain character traits influence RQ has already been addressed in preceding research. For the *Big Five*, higher actor than partner effects – as well as no, or only exceedingly slight, additional effects of partner similarities – for RQ prediction were reported: in three very large nationally representative samples of married couples from Australia, the United Kingdom, and Germany, actor effects accounted for approximately 6% of the variance in *relationship satisfaction*, while partner effects explained 1 to 3% and similarity effects less than 0.5%, respectively after controlling for actor and partner effects (Dyrenforth, Kashy, Donnellan, & Lucas, 2010).

Studies on the incremental effects of similarity, e.g., in attitudes, values, life goals, and other traits, have so far been inconsistent. In some countries, additional minor effects were found: e.g., in a large German study predicting *break-up* after one year (Becker, 2012) and in two nationally representative Chinese studies predicting *relationship satisfaction* (George, et al., 2015). In contrast, two representative Dutch studies did not find a significant additional effect of similarity (Barelds, 2005).

To conclude, contributions of different personality effects have only been estimated roughly without applying potentially more conclusive modeling methods from data-science thus far (Yarkoni & Westfall, 2016). Inherent CV, mechanisms of selecting and engineering variables could be a first step to evolve a more complex model which can improve determining the possible additive and interactive effects of actor, partner and similarity variables.

2.1.2. Trait-specific effect of similarities

Preceding research, e.g., by Keizer and Komter (2015) and George et al. (2015), showed inconsistent results when comparing the impact of different similarity scores. These studies hinted at an evident complication: it is plausible that, for varying traits, different combinations of both partner's values affect RQ diversely. For instance, another experimental study suggested that optimists profit from optimistic partners more than pessimists (Robert Böhm, 2010). Heterogeneous variable engineering for combining the scores of two partners would address this issue, but existing studies, e.g., by Keizer and Komter (2015) and Becker (Becker, 2012), used only one or two of the following score techniques at a time to calculate

---

[4] *Preceding background section which is additionally relevant for this study:*
1.1.2. Reproducible success of previous prediction models

the similarities: differences of both partner's characteristics, (profile) correlations, intra-couple correlations, scale products (moderators) or compatibility quotients. Moreover, two types of similarity scores have not yet been implemented thus far: (1) no existing score distinctly treats similarities where both partners score high, where both score low, where the actor scores higher and where the partner scores higher in a characteristic; (2) apart from a small study by Kohn (2012), preceding research has mainly focused on similarities and interactions within the same (not different) characteristics of both partners. As an example, empathy for a partner might protect them from the negative effects of the actor's insecurity on their perceived RQ.

Testing all score types in one predictive model could contribute to identifying and extracting relevant shares for different traits, as well as for different scores.

2.1.3. Effects of relationship-related and general personality

Relatively consistently across existing studies, *relationship-related personality traits,* such as attachment and love styles, have been found to be somewhat more strongly related to RQ than more *general personality traits* (Noftle & Shaver, 2006). Traits associated with a general relationship competence, i.e., as secure vs. insecure attachment style, turned out to be the most important for RQ. More *general personality traits* affected RQ only slightly: a meta-analysis (Malouff, Thorsteinsson, Schutte, Bhullar, & Rooke, 2010) as well as a cross-cultural study using representative samples from Australia, the UK and Germany (Dyrenforth, Kashy, Donnellan, & Lucas, 2010) showed that scores of four of the five-factor model personality factors correlated positively with the level of *relationship satisfaction* for the actor and the partner. The strongest associations were found for *agreeableness* and *emotional stability*, followed by *conscientiousness*, and then *extraversion*. For *openness*, results were not consistent. So far, an open research question is if general or relationship-related traits have an incremental validity for RQ prediction: they might not, because they share common variance concerning the part of personality which is relevant to social interactions.

2.1.4. Gender effects

Most available research on different characteristics whose interaction predicts RQ focused on gender differences. In a meta-analysis of the *Big Five*, inconsistent gender effects were discussed (Malouff, Thorsteinsson, Schutte, Bhullar, & Rooke, 2010). Indications for the moderated effects of gender-associated attitudes were found: e.g., two small American studies on 242 undergraduates and 289 older adults (Rudman & Phelan, 2007) resulted in

positive associations between female feministic attitudes with their own and their male partner's RQ. Moreover, the resource-related attributes of men and the beauty-related attributes of women had stronger positive effects on RQ than when the attributes than conversely (Buss, 1989). From an evolutionary perspective, these gender effects seem plausible, although differences in self-assessed partner preferences were inconsistent, small, or non-existent over different studies (Eastwick, Luchies, Finkel, & Hunt, 2014). As biologically-based moderators, these gender effects could have incremental value for predicting RQ with a model based on other relevant social, emotional, and cognitive predictors.

2.1.5. The present study

The present study is the first attempt to tackle RQ prediction based on self-assessed *personality traits* with classic methods from the ML literature. Thereby, the predictive role of similarity, actor and partner variables, different scores to scale these, as well as different personality domains was evaluated and compared. The following variables (see *Figure 3*: sets of variables, left) were used to develop (train) and cross-validate (test) the models that predicted RQ after four years (see *Figure 3*, RQ measures - right) within the longitudinal couple data of 192 partners.
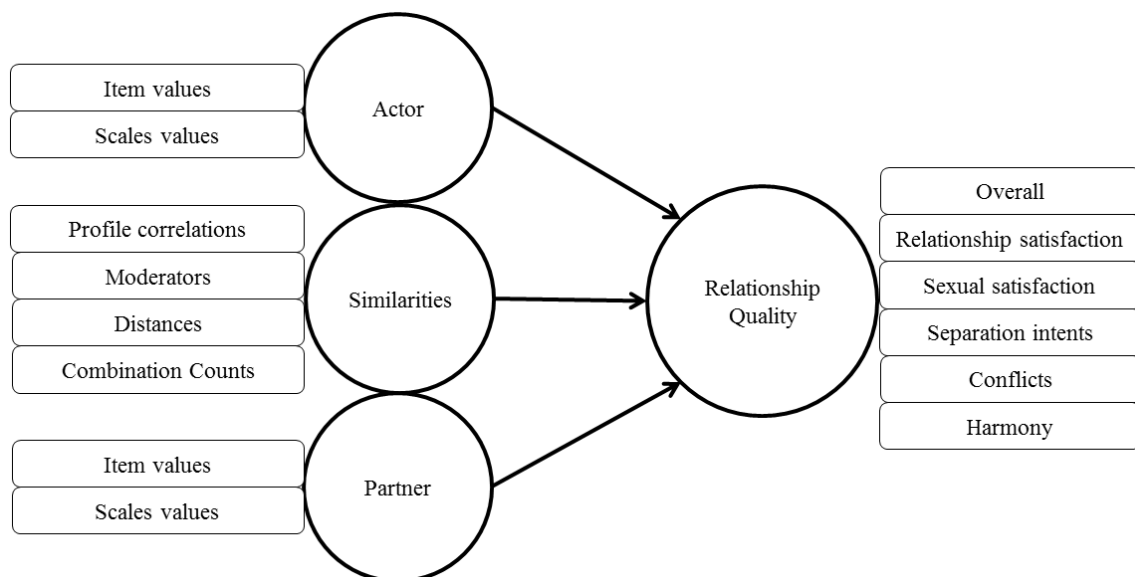
Figure 3. Linear additive model to predict different measures of *relationship quality* (right) by different sets of personality variables (left).

In a preceding assessment of predictive scale validity, models predicting *RQ overall* that were based on different scores (e.g., different methods to score partner similarity) were evaluated.

The three following hypotheses for additive linear models were evaluated:

(1) **Reproducible predictive power**

ML based models trained on all variables will be able to explain at least 25% of the variance of the overall continuous RQ, and thus out-perform simpler correlation-based approaches of former studies.

(2) **Actor-, partner- and interaction effects**

(2a) As in some prior studies using traditional regression models, in ML based models, actor, partner, and similarity variables will be able to demonstrate an incremental effect in predicting RQ beyond one another.

(2b) 245 moderators, based on hypotheses which were derived from theories and previous research, will demonstrate an incremental effect in predicting RQ. The separate moderators will only be tested each if an incremental significant effect of all combined will be found. The hypotheses mainly content gender-related effects (80). The rest (165) referred to interactions between different potential protective traits of one partner and risky traits of the other partner as e.g., social adaptiveness combined with emotional instability.[5]

(3) **Relationship-related and *general personality***

(3a) *Relationship-related personality* will demonstrate an incremental effect in predicting RQ over and beyond *general personality traits*, but not vice versa.

(3b) Models based on conflict, sex, and love-related variables, *agreeableness* and *emotional stability* will be more predictive than models with value and interest-related variables, *conscientiousness*, *extraversion* and *openness*, because they share relevant parts of the *relationship-related personality*.

In an additional explorative analysis on different RQ measures, model performances for different aspects of RQ, such as *separation intents*, *conflicts* and *relationship satisfaction*, will be evaluated.

2.2. Methods

2.2.1. Operationalization

In a longitudinal design, personality was measured at T1 and RQ was measured at T2 four years later.

---

[5] Appendix I.1. presents the complete list of moderators in the columns GGI to GPS.
  Study IV focused on hypotheses-based moderators in more detail.

2.2.1.1.   Time one personality

As is common in online dating, personality characteristics were tested with questionnaires for self-assessment (Table 3). Item formulations refer to experiences in all former close romantic relationships but never to a specific partner. Answers were given on a five-point scale: 1 as "completely false", 2 as "more false than true", 3 as "partly-partly", 4 as "more true than false", 5 as "completely true". 229 very homogeneous facet scales, mainly consisting of 5 to 10 items, were used based on the original, rationally developed scales of the PD-I (Andresen, 2015) and the BB-PI (Andresen, 2012).

To allow a differentiated variable selection, the present study analyzed many homogeneous facets instead of a few heterogeneous domains that include correlating facets. Where ever possible, facets for partner preferences, attitudes, interests and capabilities regarding the same or similar trait contents were separated into different sub-scales[6].

Table 3: Operationalization of personality variables at T1 with content domains

General Personality Personality Domain – Inventory: PD-I (Andresen, 2010) 323 items from the test construction pool	*Relationship-related personality* Bonding- and Relationship Personality – Inventory: BB-PI (Andresen, 2012) 678 items from the test construction pool
Agreeableness, emphasis on emotion and warmth	Sexuality, adventure, and desire
Pro-sociality, helpfulness, and empathy	Allurement, charm, and attractiveness
Risk appetite, thirst for adventure, sportiness	Market-orientation and pride
Neuroticism, fearfulness, insecurity	Dominance, disputability, and aggressiveness
Extraversion, gregariousness, enterprise	Unsureness, doubt, and disappointment
Conscientiousness, reliability, orderliness	Love, erotic behaviour, and understanding
Will to achieve, assignment, ambition	Troth, morals, and stability
Aggressiveness, trouble tendency, hostility	Bond, commitment, need for nearness, dependency
Openness to experiences, creative tendencies	
Intelligence, mental efficiency	
Spirit of research, will to experiment, interest in technology	

Each item and scale can be classified as

- an actor, a partner, or a similarity variable
- a *relationship-related personality* or a *general personality* variable.

  Furthermore, some of the scales can be classified as

- indicator of *emotional stability*, *extraversion*, *openness*, *agreeableness*, or *conscientiousness*
- love-related, interest-related, sex-related, conflict-related, or value-related contents.

---

[6] Appendices I.2. and I.3. present the exact allocations of items to scales and scales to domains.

2.2.1.2. Time one similarities

Similarities were calculated using the following three scores:

(1) Distances: Similarity scales were calculated by adding up the distances between the two partner's item responses. Additionally, item distances between items of both partners are added as variables.

(2) Moderators: Moderators were calculated for each scale by z-value scale partner one (actor) multiplied by z-value scale partner two. The additional hypotheses-based moderators were also combined with this score type.

(3) Combination Counts: Different combinations of item values were quantified in scores which count different combinations of actor and partner values for the same items of a scale. (Dis-)similarity combination counts emerge from combinations of low and high item values of both partners.

The overall CCs for the five-point scale with 1 as "completely false" and 5 as "completely true" were counted up as follows:

- CC_ll (low-low): counts how often item values < 3 for both partners;
- CC_hh (high-high): counts how often item values > 3 for both partners;
- CC_hl (high-low): counts how often item value actor > 3 when item value partner < 3;
- CC_lh (low-high): counts how often item partner > 3 when item value actor < 3.
  The four CCs were also calculated for every characteristic separately, e.g.:
- If both partners leveled highly in the same three of six items of own *agreeableness*, their CC_hh_A = 3;
- If both scored low in one of the other three items, their CC_ll_A = 1 (l means low);
- If the actor scored highly in the remaining two items, while the partner scored low, their CC_hl_A = 2;
- For no item did the partner score high, while the actor scores low, meaning their CC_lh_A = 0.

2.2.1.3. Time two relationship quality

Relationship stability and relationship happiness are generally considered as hallmarks of RQ (Karney & Bradbury, 1995; Harvey, Wenzel, & Sprecher, 2006; Hicks & Platt, 1970). Stability is represented by *separation intents* and actual *break-ups*. Relationship happiness is represented by perceived *relationship satisfaction*, *Sexual satisfaction*, *conflicts*, and *harmony* in different domains. The common diagnostic instruments used to measure these

aspects of RQ at T2 are described in Table 4[7]. The average of these scales was used as a measure for the general RQ (called *RQ overall).* All RQ measures were determined for each participant individually, because the perceived RQ can vary between the partners of a couple.

Table 4: Self-assessed aspects of relationship quality measures

Questionnaire	Contents (nb. of items)	Scaling	RQ measures
Questionnaire for part-nership diagnostics FDP (Hahlweg, 1996)	Amount, intensity, duration and negativity of *conflicts* (4), perceived constrictions due to current partnership (1)	1 none to 6 high	*Conflicts*
	Overall satisfaction in and with current partnership (2)	1 very dissatisfied to 7 very satisfied	*Separation intents*
Life Satisfaction Questi-onnaire FLZ (Fahrenberg, M., & Brähler, 2000)	Satisfaction with sub-aspects of life domain sexuality (7)	1 very dissatisfied to 7 very satisfied	*Sexual satis-faction*
	Satisfaction with sub-aspects sub-aspects of life domain partnership (7)		*Relationship satisfaction*
Relationship satisfaction- Inventory-Revised MSI-R (Snyder, 1997), Dyadic Adjustment Scale DAS (Spanier, 1989), Partnership Questionnaire FDP (Hahlweg, 1996)	*Harmony* in main domains within partnerships including: relationship notions (5), positive emotions (5), problem solving (5), arrangement in corporate (5), and future domains (4)	1 none to 5 high	*Harmony overall*
Marital status inventory MSI (Weiss & Cerreto, 1980)	*Separation intents* (1), *Break-up* (1, dichotomous)	Thoughts about dissolution: 0 no, 1 seldom, 2 often, 3 thoughts become intents, 4 con-crete *separation intents*, 5 serious plan to *break-up*, 6 plan already began to implement, 7 broke up	
Averages of all z-standardised above scales – which were polarised into the same direction		z-value	*RQ overall*

 Notes. *Items from MSI-R, DAS and FDP were factor analyzed with different methods, all using varimax rota-tion. The results were conflated with rational considerations: six sub-scales of harmony were derived.*

2.2.2. Sample

The sample consists of heterosexual German couples: adults, living in short or long-term relationships at T1 and having above-average educational levels.

For the 192 partners who participated at T1 and T2, the median relationship duration was Med.=41 with SD=116.5 months (Range: 1 – 519). 80 participants (41.7%) had a univer-sity degree, 61 (32.3%) had a high-school diploma (German: *Abitur*), 35 (18.2%) had fin-

---

[7] Appendix I.4. presents the RQ questionnaires and the scale construction of the harmony sub-scales in more detail.

ished secondary education and 10 (1.92%) had a lower set of qualifications, while six partici-

pants did not state their level of education. 74 (38.5%) had no children, 30 (15.6%) had one,

54 (28.1%) had two and 26 (13.5%) had more than two (maximum=6). Profile correlations of

partners for *relationship-related personality* (m=.487, SD=.165, v=.335, SE=.178, n=192)

and for *general personality* (m=.346, SD=.173, v =-.914, SE=.194, n=157) were moderate.

From 192 partners tested at T2, 55 broke up while 137 were still a couple at T2.

At T1, n=124 (64.6% of sample) are lacking less than 10% of the 4,904 personality

variables, n=22 (11.5% of sample) do not include more than 31.4% and no one is lacking

more than 54.3%. The missing values occur because only the *Stern* survey collected the

whole item pool. 436 other individuals participated at T1, but not the T2 assessment, and, as a

result, were treated as dropouts, although their personality data from T1 was used if their

partner took part in T2. Overall, the sample consists of 55 couples (110 partners) that com-

pleted T1 and T2 and 82 couples from which only one partner participated in T2.

Exclusively partners with less than 45% missing personality variables at T1 were con-

sidered for the additional cross-sectional T1-sample. Others were treated as dropouts

(n=152). *Figure 4* illustrates how many partners are lacking T1 personality variables and how

many of these variables they were deficient in. At T1, 376 of 476 subjects (79.0%) lacked

less than 1%, 57 subjects (12.0%) lacked between 6.97%, and 32.4% and 43 subjects (9.03%)

lacked between 46.6% and 54.4% of the 4,904 personality variables.

n=124 (64.6% of sample) lack less than 10%.
n=22 (11.5% of sample) lack more than 31.4%.
No sample lack more than 52.0%.

n=260 (91.5% of sample) lack less than 0.3%.
n=24 (8.45% of sample) lack between 21.7% and
54.3%. No sample lack more than 54.3%.
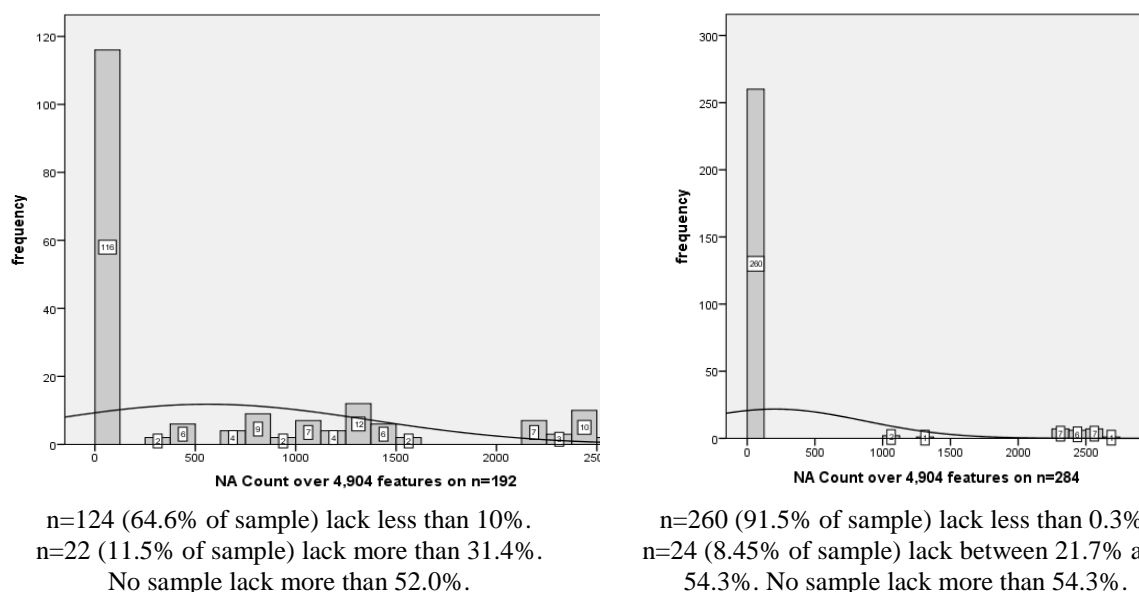
Figure 4. Missing value statistics for 4,904 personality variables at Time 1 - for partners with

Time 2 data about *relationship quality* (left) and partners who did not complete the survey at

Time 2 (right)

2.2.3. Procedure

For the ML-based evaluation, the following procedure was used. First, all variables were z-standardized. Model training and evaluation was then conducted with 10x10-fold CVs in Python[8]. To compare the relevance of certain variable groups (scales for *agreeableness*, scales for *conscientiousness* variables, actor variables, partner variables, etc.), multiple models were trained and evaluated on different sets of variables. This step was taken for the prediction of each considered RQ measure. The evaluation of the results under consideration of their static significance was examined with the corrected resampled t-test. In the following, the process and its components are described in more detail.

2.2.3.1. Elastic net regression

For the prediction of the RQ, *elastic net* regression was used. *Congruency*ing to Yarkoni & Westfall (2016) it is especially well suited to data sets with small samples and a large number of correlated variables (which exactly is the case in the present work). *Elastic net* regression optimises the weight vector w of a linear regression model

$$\hat{y} = w_0 + x_1 w_1 + ... + x_p w_p \text{, with } x_1, ..., x_p \text{ being the variable vector}$$

under consideration of two linearly combined regularization terms:

$$argmin_w(\frac{1}{2*n} * ||y - Xw||_2^2 + \alpha * \lambda * ||w||_1 + 0.5 * \alpha * (1 - \lambda) * ||w||_2^2),$$

where n is the number of samples, y is the target value vector and X is the variable matrix. Alpha is used to set the degree of regularization and lambda defines the ratio of the two regularization terms $||w||_1$ and $||w||_2^2$. The first regularization term $(||w||_1 = \sum_{j=1}^{p} |w_j|)$, which is also known as LASSO or L1 regularization, performs variable selection as well as regularization in order to improve the model's prediction performance and interpretability. The second regularization term $(||w||_2^2 = \sum_{j=1}^{p} w_j^2)$ is also known as Tikhonov or L2 regularization and aims to reduce the quadratic value of the regression coefficients, thereby shrinking large regression coefficients.

---

[8] Appendix I.5. presents the exemplary Python code.

2.2.3.2. Repeated 10-fold cross-validation

The performance of each *elastic net* model was evaluated using 10-fold CVs. There-fore, the dataset was split into 10 roughly equally sized subsamples with both partners of a couple always being in the same subsample. An *elastic net* model was then developed and optimized on 9 of these 10 folds (training set) and applied to the remaining fold (test set). This was repeated ten times, so that each fold was once used for model evaluation.

Before beginning to train the models, missing values were replaced by the mean of the non-missing values. The replacement was conducted within the CV procedure so that the mean was calculated only based on the values of the training set and then used to replace missing values of training set and test set. This ensures that no information of the test set leaks into the training set.

To enhance the validity and reproducibility of the results (Yarkoni & Westfall, 2016), the present study worked with 10 times repeated 10-fold CV, i.e. the process described was repeated 10 times, each time with different splits for the CV folds. The overall performance of the *elastic net* model is then given by the average performance of the models of the differ-ent test sets.

2.2.3.3. Parameter optimization

The selection of the hyper-parameter alpha was incorporated into the CV procedure while lambda was set fix to $\lambda = 0.5$. Before training for each model began, an inner CV was carried out on the training set of the outer CV. This process was repeated using different val-ues for alpha; the alpha value which led to the best results in the inner CV was then used for model training of the outer CV. This procedure is called nested CV. It is necessary to avoid a potential overfitting of the hyper-parameters as described by Cawley et al. (2010). In scale evaluation, a quicker hyper-parameter tuning was used only for these calculations. The so-called Grid Search adapted lambda between 0 and 1 for Alpha = 1, =.55 and .1 (Kuhn, 2016).

2.2.3.4. Evaluation measures

To evaluate the predictive power of the regression models, the mean squared error (MSE) and the coefficient of determination (r^2) were used. The MSE is given by

$$MSE = \tfrac{1}{n} \sum_{i=1}^{n} (\hat{y}_i - y_i)^2 \, ,$$

with y being a vector of n observations of the RQ and y being the corresponding predictions of the evaluated model. In the present study, all variables were z-standardized. An MSE=1

reflects deviations of predicted and observed values about one SD because all variables were z-standardized. The coefficient of determination is given by

$$r^2 = 1 - \frac{\sum\limits_{i=1}^{n} (\hat{v}_i - \bar{y})^2}{\sum\limits_{i=1}^{n} (v_i - \bar{y})^2} \quad ,$$

with y being the average of the vector y. Please note that, since model training and model evaluation are carried out on different data sets, the coefficient of determination may be negative.

The standard t-test is known to have a high type one error when used to evaluate results from repeated CV (Dietterich, 1998) due to the overlapping and the resulting dependence of the different trainings and test sets (Nadeau & Bengio, 2003). Therefore, the corrected resampled t-test was used, as it is especially suited to evaluating results generated with repeated CV (Bouckaert & Frank, 2004).

2.3. Results[9]

2.3.1. Preliminary analysis

2.3.1.1. RQ measures

Table 5 details the descriptive statistics of the RQ measures and their inter-correlations (Pearson): last mentioned were generally positive and ranging from low to high (.85> r >.15), indicating they belong to the same construct (RQ) but that they also relevantly differ from one another. RQ measures were positively correlated (.8> r >.5) between partners.

---

[9] Appendices I.6. to I.11. present the results in detail.
 Appendix I.12. presents the content of some extra models based on the simple grid search.

Table 5: Descriptive statistics and inter-correlation (Pearson) for relationship quality measures (n=192)

	Mn	SD	CA	P12 r	1	2	3	4	5	6	7	8	9	10	11	12
1 Harmony overall	3.78	.691	.944	.76												
2 Conflicts	3.38	1.16	.856	.58	.54											
3 Relationship satisfaction	5.43	1.28	.897	.65	.83	.51										
4 Sexual satisfaction	5.14	1.46	.909	.65	.50	.23**	.57									
5 Separation intents	.242	.954	.892	.52	.79	.47	.79	.26								
6 Break-up				1	.46	.30	.39	.16*	.52							
7 Relationship notions	3.87	.74	.775	.65	.86	.51	.71	.86	.69	.43						
8 Corporate domains	3.78	.76	.732	.62	.84	.46	.65	.84	.67	.46	.72					
9 Future domains	3.80	.79	.764	.62	.86	.54	.75	.86	.73	.41	.73	.74				
10 Problem solving	3.71	.76	.728	.52	.88	.46	.79	.88	.71	.34	.73	.78	.67			
11 Positive emotions	3.88	.80	.829	.79	.83	.38	.74	.83	.67	.39	.71	.70	.68	.75		
12 (Time 1) *Separation intents*	.476	.484	.821	.63	.39	.30	.37	.26	.36	.29	.28	.31	.27	.37	.29	

Notes. 1-6: Main relationship quality measures. 7-12: subscales of harmony. SD: standard deviation. CA: Cronbach's alpha.
Exceptions for sample size: For descriptive statics of separation intents at T1, n=476; r P12: Intra-couple correlation on 110 partners.
* p<.05, **p<.01, all other correlations: p<.001.

2.3.1.2. Predictive scale validity

Table 6 presents predictive scale validities with the *RQ overall* for all 223 facets and for the seven different scores (actor and partner values, distances, and the four CCs). Distance scales were correlated with the *RQ overall* exclusively negatively –if significant - showing that dissimilar partner values were worse for RQ than similar values.

Altogether, the predictive validities of the CCs were similarly high compared to the distance, actor, and partner scales. For some contents, only the CCs, but not actor or partner values, correlated with *RQ overall*, indicating independent interaction effects (e.g., mixEO139p: partner preference for *emotionality*, _LZ036a: affectionate tender attitude, mixEI159a1: inner balance, mixRE143: traveling attitude and preferences).

For the CCs, three correlation patterns were observed:

(1) **Shared weaknesses and shared strengths**

For many of the original BB-PI-scales with negatively connoted content, both partners scoring low was positively correlated, and both partners scoring high was negatively correlated with *RQ overall* (MB004: suspicious-insecurely attached, EK020: jealous-controlling, SG069: shy-inhibited, KU053: disturbed communicating-misunderstood, BH086: relationship-questioning). This finding was reversed for scales with positively connoted content (GZ049: happy-satisfied, BB093: preserving-relationships).

(2) **Negative dissimilarities and irrelevant similarities**

For some scales, only the CCs, which measure dissimilarities, negatively correlated with *RQ overall*, while CCs measuring similarities did not (TG015: loyally attached, LK041: motivated to achieve-carrier orientated, LF097: fixed on lover's qualities, IE112: intercultural-exotic loving).

(3) **Preference fittings**

For preference-related facets, the actor scoring high and the partner scoring low was found to negatively affect RQ (e.g., NV001: need for closeness and affiliation, TG015: loyally attachment).

The correlations to *RQ overall* show that the subdivided facets (e.g., preferences, attitudes, interests, capabilities) in some cases differed in their predictive validity, despite them consisting of many similar items. For instance, with reference to the three facets on top of the *Table 6*: no score, that combines preferred and own *sexual activeness* in one scale, significantly correlated with the RQ four years later. However, for the scale which only includes items of 'sexual activeness' as an attitude and excludes items for preferred 'sexual active-

ness' was shown: both partners scoring low was positively correlated to *RQ overall* (CC_ll: r(190)=.26, p<.001); the actor scoring high, while the partner scored low was negatively correlated with *RQ overall* (C_hl: r(190)=-.25, p<.001).

Table 6: Predictive scale validities for relationship quality overall after four years (n=192)

Facets related to sex		Actor effects		Partner effects		Distances		CC hh	CC ll	CC hl	CC lh
		r	N	r	N	N	r	r	r	r	r
BB-PI ST023	sexual drive (all)	.07	243	.02	190	188	-.11	.11	-.01	-.12	-.12
mix ST023p	sexually active (preference)	.01	246	.05	192	192	-.15 ·	.10	.01	-.07	-.06
mix ST023a	sexually active (attitude)	-.07	246	-.16 ·	192	192	-.15 ·	.03	.26 ···	-.13	-.25 ···
BB-PI DS014	being a rough sexist (all)	.03	246	.00	170	170	-.08	.04	.01	-.03	-.10
_BB-PI DS014p	crude sexism imaginative (preference)	.03	246	.03	170	170	-.03	.00	-.05	-.01	-.08
_BB-PI DS014t	tolerating crude sexism	.02	246	-.03	170	170	-.02	.08	.08	-.03	-.08
BB-PI RE064	erotic-imaginative	-.05	246	-.02	170	170	.03	.00	.09	-.10	-.10
_BB-PI RE064a	erotic-imaginative (attitude)	-.04	245	.02	191	190	-.08	.10	.03	-.20 ·	-.06
_BB-PI RE064p	erotic-imaginative (preference)	-.04	246	.05	192	192	-.09	.06	.02	-.07	-.03
mix EF138p	erotic-imaginative (preference, all)	-.13	246	.01	192	192	-.18 ·	.03	.12	-.06	.00
BB-PI ZE117p	tender-erotic (preference)	-.07	243	.08	190	188	-.01	-.04	.11	-.10	.02
BB-PI WS039	sexually-bold (all)	-.01	192	-.12	139	128	-.14	-.02	.07	.04	-.14
_BB-PI WS039p	sexually adventuresome (preference)	.05	192	-.05	139	128	-.12	.01	-.05	.08	-.13
_BB-PI WS039a	sexually adventuresome (attitude)	-.03	246	-.13	187	187	-.13	.12	.07	.07	-.18 ·
BB-PI PT019	primitive-instinctive (attitude)	.02	228	.05	160	150	-.05	.04	-.06	-.06	-.18 ·
_BB-PI SD107a	sexually daring (attitude)	.07	246	-.01	170	170	-.18 ·	.04	.07	-.04	-.11
BB-PI PD099	perverse-deviant (attitude)	-.06	246	-.18 ·	170	170	-.01	-.11	.08	.08	-.01
BB-PI PA074	punitive- antideviant (anti-preference)	-.03	244	-.03	191	188	-.06	-.04	-.03	.04	.03
BB-PI LB072	qualified as lover	.13 ·	246	.10	192	192	-.10	.14	.09	-.06	.00
_BB-PI LB072c	qualified as lover (capability)	.11	246	.15 ·	192	192	-.08	.13	-.06	-.12	-.01
BB-PI LF097	fixed on lover's qualities (attitude)	.09	246	.06	192	192	-.18 ·	.12	-.06	-.10	-.26 ···
BB-PI LA092	demanding lover's qualities (preference)	-.04	246	-.05	192	192	-.11	-.01	.00	-.04	-.04
mix LQ200p	lover's qualities (preference, all)	-.02	246	.04	170	170	.03	.01	-.09	-.03	-.04
BB-PI IA031	intimacy-aversive	-.15 ·	246	.01	192	192	-.12	-.02	.15 ·	-.16 ·	-.01
BB-PI SV084	fearing sexual failure	-.12	242	-.05	189	186	-.09	-.15 ·	.12	-.12	-.05
BB-PI SU002	sexually unsatisfied	-.31 ···	246	-.39 ···	187	187	-.28 ···	-.15	.47 ···	-.10	-.29 ···

Facets related to *extraversion*, *openness*, *conscientiousness*, values and interests	Actor effects r	N	Partner effects r	N	Distances N	r	CC hh r	CC ll r	CC hl r	CC lh r
BB-PI PB025 phlegmatic-comfortable (all)	-.11	246	-.27 ***	192	192	-.25 ***	-.12	.30 ***	-.08	-.20 ***
mix PB025p phlegmatic-comfortable (preference)	.08	246	.02	192	192	.01	.09	-.01	.01	-.05
_BB-PI PB025a phlegmatic-comfortable (attitude)	-.11	246	-.27 ***	192	192	-.25 ***	-.12	.30 ***	-.08	-.20 ***
BB-PI RK032 talkative-communicative (all)	-.08	246	-.11	187	180	-.15 *	-.01	.12	-.02	-.03
_BB-PI CO168p communicative (preference)	-.04	246	-.01	170	170	-.25 ***	.06	.10	-.06	-.07
_BB-PI RK032p talkative-communicative, need (preference)	.10	245	.10	192	192	-.13	.20 ***	-.02	-.14	-.12
BB-PI SA007 beauty-conscious – attractive (capability)	-.08	246	.02	192	192	-.10	.08	.03	-.13	-.06
BB-PI SH096 in love with beauty-hyper aesthetic (preference)	.00	246	.08	187	187	-.09	.14	.05	-.12	.05
BB-PI SF089 eager to look good (attitude)	-.08	246	.02	170	170	.02	.07	.03	-.05	.10
BB-PI CG102 charming-winning (capability)	.09	246	.00	192	192	-.06	.15 *	-.12	.14	-.05
mix CG102a sexy charming (attitude)	.02	211	-.12	152	184	-.15 *	.13	.19 *	-.01	-.18 *
mix CG102a spirited (attitude)	.11	241	-.02	187	182	-.05	.07	-.04	.09	-.13
mix CG102 distanced-shy (attitude)	-.23 ***	246	-.06	170	170	-.24 ***	-.15	.27 ***	-.26 ***	.01
BB-PI SG065 shy-inhibited (attitude)	-.32 ***	246	-.08	170	170	-.13	-.22 ***	.25 ***	-.33 ***	.02
PD-I CG10c charming (capability)	.11	241	.02	165	160	-.05	.14	.04	.04	-.10
mix CH127c charismatic (capability)	.13 *	246	.08	192	192	-.10	.13	-.18 *	-.09	-.03
BB-PI HG037 histrionic-compliant (attitude)	-.10	192	-.06	139	128	-.01	-.05	.14	-.13	-.02
mix AU128c radiating (capability)	.16 *	210	-.02	151	192	-.05	.17 *	-.09	.00	-.06
PD-I TE126a spirited (attitude)	.07	241	-.01	165	160	-.01	.07	-.01	.02	-.11
mix UL130p enterprising (preference)	-.01	246	-.01	192	192	-.07	-.06	-.02	.04	-.06
PD-I UK075a enterprising (attitude)	-.01	241	-.04	165	160	-.08	-.04	.02	.08	-.06
BB-PI UK075 enterprising-outgoing	-.08	246	-.12	192	192	-.15 *	-.10	.09	-.01	-.11
_BB-PI UK075p enterprising-outgoing (preference)	-.03	246	-.06	192	192	-.15 *	-.02	.07	.03	-.13
_BB-PI UK075a enterprising-outgoing with(out) partner (attitude)	-.07	246	-.11	192	192	-.17 *	-.10	.14	-.04	-.07
PD-I PG125i partying and sociability (interest)	.08	241	.01	165	160	-.17 *	.22 *	-.01	.00	-.17 *
mix PG125p partying and sociability (preference)	-.05	246	-.06	192	192	-.19 *	.00	.08	-.04	-.15 *
PD-I GE129p sociability (preference)	.12	241	.00	165	160	-.04	.10	.03	.08	-.10
PD-I GE129- reserved (attitude)	-.19 ***	241	-.02	165	160	-.19 *	.00	.25 ***	-.19 *	-.04
PD-I GE129all sociable (all)	.16 *	188	-.04	139	160	-.04	.10	.03	.08	-.10
BB-PI AS079 adventurous-sporty (all)	-.11	245	-.13	191	190	-.14	-.14	.19 *	-.03	-.05
_BB-PI AS079p adventure seeking (preference)	.02	241	-.04	165	160	-.18 *	.23 ***	.11	-.12	-.17 *
_BB-PI AS079p sportive (preference)	-.06	241	-.10	165	160	-.16 *	-.05	.10	-.07	-.19 *
PD-I AS079p experience seeking (preference)	.03	241	-.05	165	160	-.18 *	-.03	.09	-.08	-.25 ***
PD-I SP166i sportive (interest)	-.06	246	-.06	170	170	-.16 *	-.05	.05	-.10	-.20 *
mix SS142p sensation seeking (preference)	-.08	246	-.13	170	170	-.13	.03	.10	-.08	-.18 *
PD-I EN144a exploring (attitude)	.01	241	-.13	165	160	-.17 *	-.09	.10	-.04	-.20 *

Facets related to *extraversion*, *openness*, *conscientiousness*, values and interests		Actor effects		Partner effects		Distances		CC hh	CC ll	CC hl	CC lh
		r	N	r	N	N	r	r	r	r	r
BB-PI HD063	domineering-dominant (attitude)	-.11	246	-.10	192	192	-.09	.08	.20 ·	-.08	-.01
BB-PI FD055	leadership quality-dominant (all)	-.05	246	-.08	192	192	.01	-.08	.05	.04	.01
_BB-PI FD055c	commanding (capability)	.08	243	.01	191	188	.01	.06	-.05	.02	-.02
mix DO164p2	dominant (preference)	-.05	246	-.02	192	192	-.15 ·	.08	.12	-.13	-.13
_BB-PI DO164c	protecting (capability)	.24 ***	246	.09	192	192	.02	.19 ·	-.11	-.05	-.03
mix DO164p	dominant & protective (preference)	-.19 ***	246	.06	192	192	-.08	.07	.07	-.12	.02
mix DO164p1	Protective (preference)	-.23 ***	246	.12	192	192	-.06	.04	.00	-.10	.12
BB-PI BV105	protective-defending (attitude)	.24 ***	246	.05	192	192	.01	.15 ·	-.19 ·	.00	-.08
BB-PI VS114	providing-securing	.07	231	.08	186	178	-.11	.17 ·	-.04	-.07	-.05
BB-PI MG110	materially generous	-.01	246	-.06	192	192	.08	-.05	.03	.03	.04
BB-PI GA038	emotionless-alexithym (all)	-.27 ***	246	-.22 ***	170	170	-.16 ·	-.03	.33 ***	-.19 ·	-.10
_BB-PI GA038p	emotionless-alexithym (preference)	-.31 ***	214	-.11	157	148	-.25 ***	.03	.34 ***	-.32 ***	.01
_BB-PI GA038a	emotionless-alexithym (attitude)	-.20 ***	246	-.23 ***	170	170	-.13	-.06	.24 ***	-.10	-.12
PD-I KR160c2	having a free mind/ fantasy/ideas (capability)	-.01	241	-.06	165	160	-.14	.07	.04	-.01	-.16 ·
PD-I KR160c	creativity (capability)	-.02	241	-.03	165	160	-.17 ·	.07	.04	-.06	-.19 ·
PD-I KR160p	creativity (preference)	.06	246	.03	192	192	-.02	.08	-.01	.01	-.13
PD-I KR160p-	against creativity (preference)	.02	241	.03	165	160	-.02	-.01	-.02	-.05	.04
mix EH136p	achievement striving (preference, all)	-.02	246	-.12	192	192	-.13	-.10	.05	.00	-.16 ·
mix EH136p2	support for achievement striving (preference)	.01	242	-.16 ·	189	186	-.10	-.13	.07	.03	-.14
PD-I EH136a	achievement striving (attitude)	-.02	241	.01	165	160	.01	.02	.00	-.01	-.05
PD-I LE137c	performance capacity (capability)	.07	241	.04	165	160	-.01	.12	-	-.11	-.20 ·
mix AP137ri	work priority (attitude)	-.08	245	-.16 ·	191	190	-.16 ·	.06	.19 ·	-.12	-.16 ·
BB-PI WA066	workaholic (attitude)	-.11	246	-.15 ·	192	192	-.17 ·	.05	.22 ***	-.16 ·	-.05
_BB-PI LK041a	carrier oriented: lack of time for partner (attitude)	.03	242	-.09	190	188	-.19 ·	.05	.06	.00	-.25 ***
BB-PI LK041	carrier orientated: motivated to achieve (attitude)	.03	243	-.15 ·	190	188	-.09	-.11	.06	.03	-.23 ***
MK021	artistic-cultural (preference for)	.00	246	-.05	192	192	-.03	.00	-.01	.06	-.11
PD-I MK021c	artistic-cultural (capability)	.00	241	-.04	165	160	-.21 ·	.04	.04	-.02	-.14
BB-PI ES018	esoteric-spiritual (attitude)	.03	242	.04	189	186	-.06	.06	-.02	-.03	-.03
PD-I ES018i	spiritual (interest)	-.03	241	-.07	165	160	-.11	.00	.10	-.04	-.11
_BB-PI ES018p	esoteric-spiritual in partnership (preference)	-.02	242	.01	189	186	-.08	.05	-.01	-.03	-.13
mix ES018a	esoteric-spiritual (all, attitude)	-.01	245	.06	191	190	-.07	.06	-.08	-.03	-.01
mix SP162i	spiritual in partnership (interest)	.05	245	.06	191	190	.00	.07	-.02	-.02	.04
PD-I SP162i	spiritual (all, interest)	-.05	241	-.07	165	160	.00	-.09	.04	-.06	-.01
PD-I TE167i	technologic (interest)	-.09	241	-.12	165	160	.05	-.17 ·	.18 ·	.00	-.05
_BB-PI LI006i	love topics (interest)	-.22 ***	242	.01	189	186	-.16 ·	-.01	.14	-.09	-.05
mix RE124	religious	-.06	246	-.17 ·	192	192	-.05	-.18 ·	.23 ***	-.07	-.05

Facets related to *extraversion*, *openness*, *conscientiousness*, values and interests		Actor effects		Partner effects		Distances		CC hh	CC ll	CC hl	CC lh
		r	N	r	N	N	r	r	r	r	r
BB-PI KW080	church-traditional valuing	-.05	246	-.09	192	192	-.10	-.12	.18 ·	-.06	.01
PD-I RE081p	wild traveling (preference)	-.06	241	-.10	165	160	-.08	-.04	.16	.02	-.11
PD-I RE081p2	conventional traveling (preference)	.06	241	.10	165	160	.08	.17 ·	-.10	-.11	-.15
mix RE143	traveling (all)	-.07	245	-.12	191	190	-.21 ***	-.08	.26 ***	-.10	-.08
mix RE143c	traveling (capability)	-.09	244	-.12	170	170	.04	-.20 ·	.14	.15	-.04
BB-PI IE112	intercultural-exotic loving	.02	242	-.07	189	186	-.09	-.03	.07	.00	-.23 ***
_BB-PI CU165p	partner of other culture (preference)	-.01	246	-.03	192	192	-.07	.02	.04	-.05	-.05
mix KF013a	children and family person (attitude)	.02	246	-.02	192	192	.08	.07	.03	.04	-.02
mix KF013p	children and family person (preference)	.02	246	.18 ·	170	170	-.01	.21 ·	-.08	-.02	-.04
mix FA013	sense for family	.04	246	.13	170	170	.05	.17 ·	-.07	.07	.01
mix KI013	fond of children	.01	246	.04	192	192	.08	.08	.05	.00	-.06
mix KI013i	children and family (all, interest)	.02	246	.05	192	192	.09	.14	.03	.00	-.04
BB-PI GE090	demanding *conscientiousness*	-.07	246	.00	192	192	-.09	-.07	.16 ·	-.02	-.05
mix CO163p	conscientious (preference, all)	.03	246	.26 ***	192	192	-.20 ·	.01	-.02	-.08	.10
mix CO163p2	*manners* (preference)	-.01	246	.20 ***	192	192	-.04	.06	.06	-.06	-.09
PD-I CO163p3	planning (preference)	-.08	241	.09	165	160	-.15	.11	.03	-.17 ·	.03
mix CO163a1	conscientious (attitude)	-.06	245	.07	191	190	-.17 ·	.02	.13	-.09	.02
PD-I CO163a2	regularity (attitude)	-.03	241	.16 ·	165	160	-.28 ***	.12	.10	-.37 ***	.00
BB-PI ZS050	ensuring future	-.03	244	-.03	191	188	-.17 ·	.03	.03	-.05	-.09
BB-PI GF106	habit fixed	.02	228	.05	185	178	-.09	.12	.15 ·	-.12	.02
PD-I IQ158c1	logical reasoning (capability)	-.06	241	.02	165	160	-.08	-.07	-.01	-.13	.08
PD-I IQ158c2	intelligent (capability)	.09	241	.11	165	160	-.03	.08	-.01	-.03	.03
PD-I IQ158c3	mentally speedy (capability)	.03	241	.02	165	160	.09	.05	.11	-.13	-.14
PD-I IQ158c4	skilled in problem solving (capability)	.01	241	.04	165	160	.08	-.08	.10	-.10	-.04
PD-I IQ158c5	capable to learn (capability)	.20 ***	241	.08	165	160	-.09	.11	.00	-.08	-.10
PD-I IQ158c6	ingenious (capability)	.08	241	.02	165	160	.04	.04	.08	-.04	-.03

Other facets related to love practice		Actor effects r	N	Partner effects r	N	Distances N	r	CC hh r	CC ll r	CC hl r	CC lh r
BB-PI TG015	loyally attached	-.05	226	.14	183	174	-.21 ***	.13	.01	-.20 *	-.04
mix TR141a	faithful (attitude)	.03	246	.27 ***	192	192	-.39 ***	.19 *	.05	-.18 *	-.03
mix TR141r	religious faithfulness	-.08	246	-.01	192	192	-.05	-.08	.10	-.06	.08
mix TG015a	faithfully attached (attitude)	-.02	229	.22 ***	184	174	-.32 ***	.18 *	.06	-.16 *	-.04
BB-PI PU022	polygamous-unfaithful	-.08	246	-.27 ***	187	180	-.33 ***	.06	.28 ***	-.06	-.14
_BB-PI PU022p	polygamous (preference)	-.02	246	-.10	170	170	-.22 ***	.02	.16 .	-.11	-.16 *
_BB-PI TR141-	infidelity	-.08	246	-.27 ***	187	180	-.33 ***	.06	.28 ***	-.06	-.14
_BB-PI TR141	conventional in relationships	-.01	246	-.10	170	170	-.16 *	.02	.13	-.06	-.11
BB-PI UF108	unconventional-liberal	-.02	246	-.11	170	170	-.16 *	.01	.14	-.03	-.12
BB-PI PS083	promiscuous-partner changing	-.11	245	-.11	191	190	-.09	-.04	.12	-.03	-.03
BB-PI NV001	in need for closeness and affiliation (all)	-.14 *	231	.02	186	178	-.30 ***	.11	.17 *	-.21 ***	-.10
mix NB147p	closeness (preference)	.07	212	-.04	171	162	-.22 ***	.12	.19 .	-.10	-.18 *
BB-PI DB012	distance in a partnership (preference)	-.12	242	-.02	189	186	-.27 ***	.04	.20 .	-.21 ***	-.13
_BB-PI DB012c	distance in a partnership (capability)	-.10	241	.00	189	186	-.27 ***	.08	.17 .	-.18 *	-.05
mix DB012p	distance in a partnership (preference, all)	-.10	245	.03	191	190	-.22 ***	.02	.13	-.13	-.15 *
_BB-PI DI145p	distance in a partnership (preference)	-.07	246	.03	192	192	-.29 ***	.08	.13	-.20 *	-.11
_BB-PI UA146p	Independent of partner (preference)	-.15 *	245	.00	191	190	-.07	.06	.12	-.06	-.01
BB-PI US035	independent- self-determined	-.20 ***	245	-.10	191	190	.01	.07	.23 ***	-.03	-.01
BB-PI UA005	involuntarily adapted	-.29 ***	245	-.13	191	190	-.17 *	.02	.28 ***	-.21 ***	-.05
_BB-PI AG161c1	adaptive (capability)	-.31 ***	246	-.16 *	192	192	-.17 *	-.07	.27 ***	-.15 *	-.07
_BB-PI UA005ri	involuntarily adaptive- risky	-.31 ***	245	-.15 *	191	190	-.13	.03	.28 ***	-.19 *	.01
_BB-PI UA005pr	functionally adaptive - protective	-.32 ***	246	-.17 *	192	192	-.21 ***	-.07	.28 ***	-.15 *	-.10
BB-PI ER009	emancipatory-radical	.05	242	.19 *	189	186	-.04	.17 *	-.06	-.08	.09
BB-PI TR024	traditional-fixed on roles	-.06	246	-.17 *	187	180	-.01	.01	.14	-.01	-.06
BB-PI MP027	macho-boastful	.01	246	.00	192	192	-.09	.05	.06	-.08	-.06
BB-PI IV008	idealizing-deifying	-.03	230	-.07	185	178	.09	-.13	.04	.09	.10
BB-PI RL094	romantic-love illusionary	-.03	245	-.10	191	190	-.18 *	.04	.17 .	-.07	-.08
BB-PI OF095	obsessive- partner fixed	-.12	246	-.13	192	190	-.12	-.08	.17 .	-.02	-.13
mix SES144p	socio-economic status (preference)	-.08	246	-.04	192	192	-.11	-.12	.08	-.11	-.02
BB-PI MO088	market value oriented	-.07	228	-.02	185	176	-.02	.02	.11	-.09	.00
BB-PI AO069	demanding-optimizing	-.12	240	-.23 ***	186	180	.01	-.04	.24 ***	.06	.03
BB-PI BS113	possessive-proud	-.06	230	-.06	186	178	-.09	-.11	.12	-.09	-.03
BB-PI AE115	demanding attention	.00	246	.18 *	170	170	-.22 ***	.18 *	.00	-.10	.00
BB-PI WI070	daring-imposing	.01	246	-.04	170	170	.04	.13	.01	.05	.02
BB-PI PV098	varying in partner type	-.06	242	-.04	189	186	-.12	.00	.12	-.01	.00

Facets related to *emotional stability*, *agreeableness* and conflict		Actor effects		Partner effects		Distances		CC hh	CC ll	CC hl	CC lh
		r	N	r	N	N	r	r	r	r	r
PD-I MS085	lack of self confidence	-.30 ***	241	-.19 *	165	160	-.19 *	-.08	.40 ***	-.23 ***	-.15
BB-PI SS028	self-damaging	-.42 ***	244	-.34 ***	187	177	-.11	-.09	.44 ***	-.14	-.22 **
mix EI159a1	inner balance	-.11	246	-.07	170	170	-.26 ***	.33 ***	.11	-.16 *	-.14
mix EI159.4	anxious	-.19 ***	241	-.13	165	160	-.27 ***	-.02	.21 .	-.14	-.16
BB-PI SN058	stressed-nervous	-.21 ***	246	-.20 *	170	170	-.22 ***	-.03	.20 .	-.08	-.12
mix EI159a3	stress robust	.11	241	.00	165	187	-.11	.20 *	-.04	.02	-.07
mix EI159.3	sensitive to critique	-.28 ***	246	-.12	187	181	-.17 *	-.16 *	.31 ***	-.19 *	-.06
mix EI159.1	pessimism	-.31 ***	246	-.29 ***	187	181	-.09	-.24 ***	.31 ***	-.09	-.13
BB-PI DN082	depressive-negativistic	-.47 ***	227	-.28 ***	160	150	-.37 ***	.02	.37 ***	-.35 ***	-.17 *
mix EI159.2	dependent	-.14 *	246	-.14	192	192	.01	.06	.14	.02	-.12
BB-PI DH073	dependent-helpless	-.06	246	-.12	192	192	-.05	.03	.11	-.05	-.16 *
BB-PI EV056	lonely-desperate	-.04	227	-.06	185	176	-.03	.04	.16 .	-.05	-.09
BB-PI FG034	fearful-security searching	-.30 ***	244	-.04	191	188	-.06	-.14	.22 ***	-.13	.11
mix EI159.5	emotional	-.21 ***	242	-.12	166	162	-.15	-.14	.28 ***	-.12	-.14
BB-PI SE044	sensible-emotional	-.36 ***	246	-.21 *	170	170	-.11	-.13	.37 ***	-.19 *	-.12
BB-PI LE060	passionate-emotional	-.24 ***	246	-.13	170	170	.06	-.05	.22 ***	-.06	-.10
BB-PI EK020	jealous-controlling (all)	-.23 ***	242	-.14	189	186	-.10	-.17 *	.25 ***	-.11	-.03
_BB-PI EK020a	controlling when jealous (attitude)	-.17 *	242	-.13	189	186	-.21 ***	-.05	.25 ***	-.12	-.07
mix EK020p	jealous (preference)	-.22 ***	243	-.11	190	186	-.06	-.21 ***	.18 .	-.12	.00
BB-PI VV101	willing to reconcile-forgive (all)	.17 *	245	.11	187	181	-.01	.12	-.08	-.04	.00
_BB-PI VV101a	forgiving (attitude)	.10	220	.12	177	161	-.08	.16	-.07	-.07	.04
mix VV101c	forgiving & reconciliation (capability)	.20 ***	246	.15	170	170	-.06	.24 ***	-.09	-.12	-.24 ***
mix AG161c2	forgiving (capability)	.14 *	246	.13	170	170	-.10	.17 *	.00	-.14	-.23 ***
BB-PI GH103	good-natured-humorous–agreeable (all)	.14 *	246	.31 ***	170	170	-.20 *	.28 ***	-.11	-.20 *	-.11
mix GH103c	humorful (capability)	.23 ***	246	.21 *	170	170	.05	.25 ***	.06	-.10	-.12
mix AG161p	agreeable (preference)	.09	246	.19	112	112	-.13	.11		-.01	.05
mix UM131c	accessible (capability)	.27 ***	246	.33 ***	192	192	-.07	.35 ***	-.12	-.16 *	-.17 *
BB-PI TA051	tolerant-accepting (attitude)	.09	245	.12	192	192	-.06	.16 *	-.07	-.10	-.09
mix HA132p	harmonic (preference)	.05	246	.19 .	170	170	-.19 *	.22 ***	-.08	-.16 *	-.08
BB-PI VE111	understanding-sensitive (all)	.20 ***	246	.19 *	192	192	-.23 ***	.37 ***	.07	-.21 ***	-.16 *
_BB-PI VE111c	understanding-sensitive (capability)	.24 ***	246	.21 **	192	192	-.24 ***	.38 ***	.07	-.24 ***	-.17 *
_BB-PI VE111p	understanding-sensitive (preference)	.06	191	.12	139	128	-.16	.14		-.10	.05
mix ZU054c	listening (capability)	.21 ***	246	.15 .	170	170	-.19 *	.21 *	-.06	-.11	-.18 *
mix EO139p	emotional (preference)	-.01	246	.10	192	192	-.33 ***	.20 ***	.00	-.19 *	-.09
_BB-PI LZ036a	affectionate tender (attitude)	-.01	245	.08	191	190	-.27 ***	.17 *	-.01	-.19 *	-.01

Facets related to *emotional stability*, *agreeableness* and conflict		Actor effects r	N	Partner effects r	N	Distances N	r	CC hh r	CC ll r	CC hl r	CC lh r
PD-I EM133c	empathic (capability)	.11	241	.18 *	165	160	-.13	.26 ***	-.06	-.14	-.18 *
BB-PI FH062	caring-helpful	.08	242	.24 ***	189	186	-.15 *	.19 *	-.02	-.20 *	-.05
_BB-PI SS134c	social supportive in partnerships (capability)	.16 *	245	.27 ***	191	190	-.27 ***	.26 ***	.00	-.23 ***	-.09
PD-I SS161_BB-PI c2	socially supportive (capability)	.09	241	.07	165	160	-.11	.16 *	-.12	-.06	-.14
PD-I SS161_BB-PI c1	socially engaging (capability)	.06	241	.01	165	160	.05	.06	-.06	.04	.00
_BB-PI HE043a	enforcing ego (attitude)	-.06	246	.02	192	192	.00	.05	.05	.02	.01
BB-PI HE043	hard-egoistic	-.18 ***	246	-.28 ***	170	170	-.24 ***	-.06	.27 ***	-.12	-.21 *
mix EI159a2	egoistic (attitude)	-.25 ***	246	-.21 ***	192	192	-.17 *	-.20 *	.33 ***	-.19 *	-.13
_BB-PI AG161c	stubborn (incapability)	-.26 ***	246	-.36 ***	170	170	-.27 ***	-.13	.45 ***	-.16 *	-.27 ***
mix DI135a	pigheaded (attitude)	-.07	175	-.20 *	131	162	-.26 ***	-.07	.37 ***	-.07	-.19 *
BB-PI EA052	arrogant-overweening	-.13 *	246	-.12	192	192	-.20 *	.27 ***	.20 .	-.11	-.15 *
BB-PI BE109	insulted-huffy	-.28 ***	246	-.34 ***	170	170	-.22 ***	-.13	.45 ***	-.21 *	-.20 *
BB-PI RN068	vindictive-vile	-.10	246	-.05	192	192	-.16 *	.00	.17 .	-.17 *	-.03
PD-I AG161.1	vindictive	-.07	241	-.17 *	165	160	-.13	-.08	.21 .	.02	-.13
BB-PI UG061	weary-bored	-.37 ***	245	-.32 ***	170	170	-.23 ***	-.07	.33 ***	-.19 *	-.19 *
BB-PI NA046	negativistic-pejorative	-.57 ***	246	-.19 *	170	170	-.19 *	.03	.27 ***	-.27 ***	.02
BB-PI SZ047	pugnacious-angry	-.21 ***	230	-.26 ***	186	178	-.07	-.12	.24 ***	-.01	-.01
BB-PI KA011	physically abusive	-.17 *	246	-.30 ***	192	192	-.20 *	-.05	.28 ***	-.10	-.27 ***
BB-PI VA016	verbally aggressive	-.12	205	-.29 ***	171	162	-.17 *	-.05	.22 ***	-.02	-.25 ***
BB-PI KU053	disturbed communicating-misunderstood (all)	-.37 ***	228	-.43 ***	179	160	-.22 ***	-.32 ***	.49 ***	-.13	-.23 ***
_BB-PI KU053ac	active -misunderstanding	-.24 ***	214	-.24 ***	157	148	-.13	-.06	.34 ***	-.02	-.09
_BB-PI KU053pa	passive -misunderstood	-.37 ***	228	-.44 ***	179	160	-.22 ***	-.36 ***	.48 ***	-.09	-.23 ***
BB-PI KV040	avoiding *conflicts*-closed (all)	-.23 ***	245	-.18 *	170	170	-.16 *	-.06	.26 ***	-.14	-.10
_BB-PI KV040ri	remaining silent as a method	-.33 ***	215	-.25 ***	157	148	-.12	-.10	.30 ***	-.15	-.07
_BB-PI CO168a-	avoiding *conflicts*	-.18 ***	246	-.13	170	170	-.14	-.05	.23 ***	-.11	-.09
_BB-PI BU042	unwilling to bond/commit	-.28 ***	246	-.11	170	170	-.16 *	***	.20 .	-.14	-.02
BB-PI GZ049	happy-satisfied	.57 ***	226	.25 ***	183	174	-.24 ***	.53 ***	-.23 ***	-.04	-.32 ***
BB-PI BH086	relationship questioning	-.43 ***	246	-.38 ***	170	170	-.21 *	-.19 *	.45 ***	-.11	-.20 *
BB-PI BB093	preserving relationships	-.23 ***	246	-.15 *	192	192	-.11	-.18 *	.22 ***	-.11	-.07
BB-PI MB004	suspicious-insecurely attached	-.54 ***	216	-.28 ***	157	148	-.30 ***	-.17 *	.52 ***	-.42 ***	-.12
BB-PI VG026	trustful-salvaged	.36 ***	246	.18 *	192	192	-.12	.30 ***	.05	-.02	-.14

Notes. Pearson correlations of facets and their modifications to relationship quality overall, with levels of significance.

CC: Combination Count. hh: high-high. ll: low-low. hl: high-low. lh: low-high. In scale names: p: partner preference. c: capability. a: attitude.

PD-I: Personality Domain Inventory (Andresen, 2015). BB-PI: Attachment- and Relationship-related Personality Inventory (Andresen, 2012). mix: scales generated from items from PD-I and BB-PI. "_" prefix means that the scale deviates from the one from the original facet version.

*p<.05. **p<.01.***p<.001.

2.3.1.3. Score evaluation

To give a first overview and impression, different scores (items, scales, CCs, etc.) were vaguely pre-evaluated in their predictive validity by ML using a rough grid search. *Figure 17* presents the predictive performance of these rough *elastic net* models, based on different sets of variables.
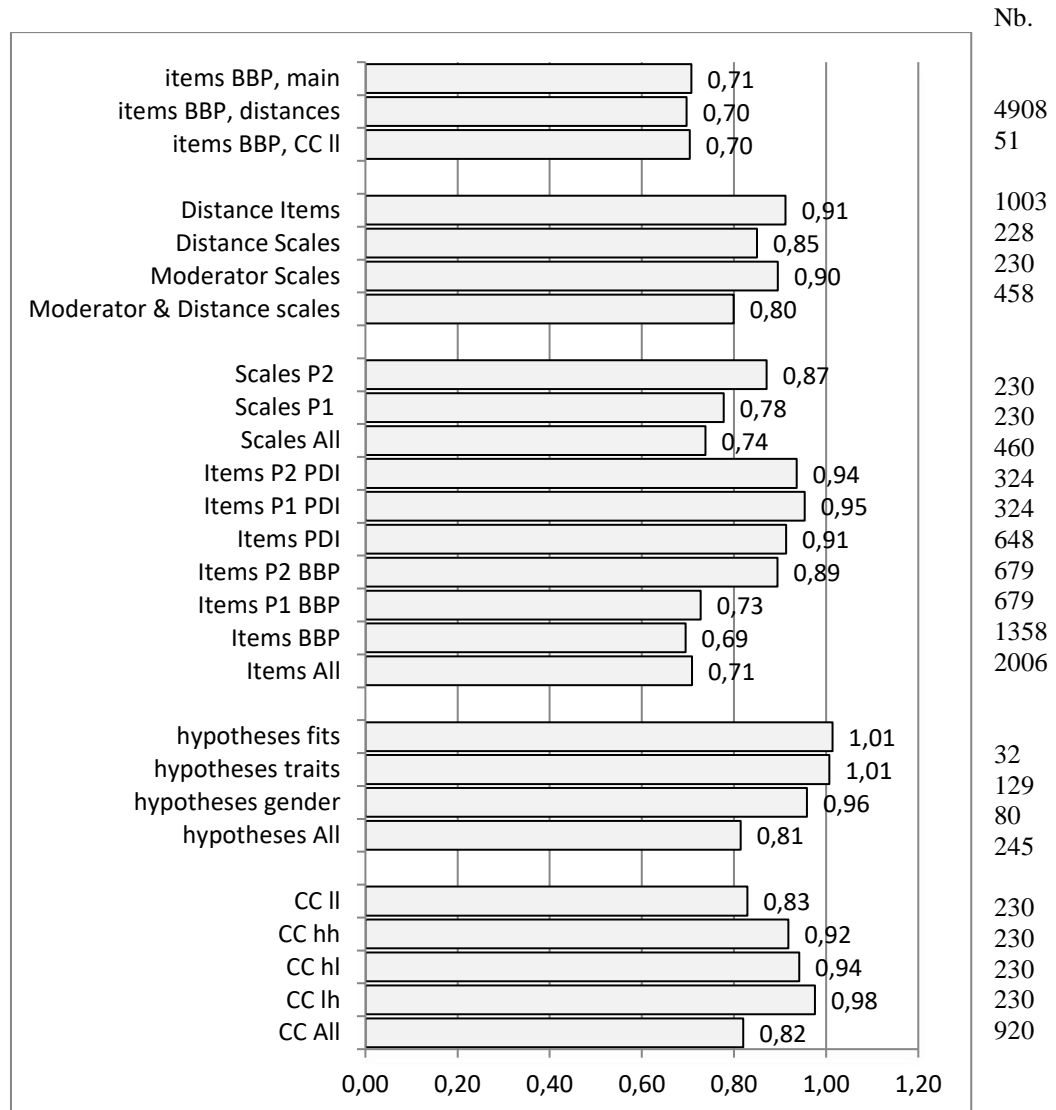
Figure 5. 10x10-fold *Cross-validation-* MSE of the *elastic net* models based on different variable sets predicting the z-value of *relationship quality* overall (n=192).

Notes. MSE: mean squared error as deviations between the values which were predicted by the model developed on the train data and the real observed values of the test data; an MSE=1 reflects a deviation of one SD because all variables were z-standardized. h: high, l: low.
BB-PI: Relationship- and Attachment-related Personality Inventory (Andresen, 2012).
PD-I: Personality Domain Inventory (Andresen, 2015).

The models that best predicted the *RQ overall* were based on all variables (MSE=.71), or on items of the BB-PI (MSE=.69). For actor variables, items (MSE=.73) compared to

scales (MSE=.74) predicted similarly accurate. Moreover, the different scores for similarities and hypotheses-based moderators predicted more accurately when they were combined in one model. To conclude, the different score types for interaction effects which combine actor and partner characteristics added up to one another incrementally. Partner, PD-I variables, and most of the CCs did not relevantly predict RQ on their own (MSE>.87): the only exception was the set of CC_ll representing the 'shared' weaknesses (MSE=.83).

2.3.2. Model performances

As described in the procedure section, a resampled CV set-up was used in combination with an appropriately modified t-test to ensure that the results are reproducible and valid despite the small sample. The reporting of confidence or credibility intervals was omitted because they are not suited to a proper evaluation of results based on repeated CV (Vanwinckelen & Blockeel, 2012). For comparison, a baseline was defined, produced by the performance of a model which always predicted the average value of the RQ measure.

Table 7 and *Table 8* present the predictive performance of models using different combinations of actor, partner, similarity, and personality domain variables with the more complex parameter tuning. It should be noted that, since model training and model evaluation are carried out on different data sets, r² may become negative.

Table 7: 10x10-fold cross-validation-performance of the elastic net models predicting main measures of relationship quality after four years (n=192)

nb.	Predictor variables	Relationship quality Overall		Separation intents		Relationship satisfaction		Sexual satisfaction		Conflicts		Harmony overall		Time 1 n=476: Separation intents	
		MSE	r²	MSE	r²	MSE	r²	MSE	r²	MSE	r²	MSE	r²	MSE	r²
	baseline	1.00	-.12	.92	-.14	1.00	-.13	1.03	-.11	1.04	-.11	.96	-.12	.88	-.05
2665	Sim, Mod	.84*	.03	.84*	-.05	.88*	.00	.73***	.23	.96	-.09	.85*	-.02	.86	-.04
245	Mod	1.05	-.20	.95	-.19	1.03	-.18	1.04	-.13	1.10	-.18	1.00	-.19	.93	-.11
2420	Sim	.81**	.07	.82*	-.03	.86*	.02	.73***	.22	.95	-.07	.80**	.03	.83	.00
5149	P1, P2, Sim, Mod	.55***	.37	.67***	.16	.69***	.21	.72***	.24	.80***	.01	.60***	.28	.85	-.03
4904	P1, P2, Sim	.55***	.37	.68***	.14	.69***	.21	.71***	.24	.80**	-.01	.60***	.28	.84	-.02
2484	P1, P2	.55***	.37	.64***	.22	.64***	.27	.69***	.27	.75***	.03	.58***	.30	.83	.01
1242	P1	.60***	.33	.66***	.19	.64***	.27	.72***	.24	.74***	-.08	.58***	.31	.83	-.01
1242	P2	.82**	.07	.84*	-.04	.88*	.00	.89*	.05	.94	.04	.82*	.01	.84	-.01
4423	R. pers., G. pers.	.55***	.37	.67***	.15	.71***	.19	.75***	.20	.80**	.02	.59***	.29	.74**	.11
3177	R. pers.	.54***	.38	.66***	.18	.68***	.21	.75***	.20	.78***	.08	.58***	.31	.90	-.08
1246	G. pers.	.88*	.01	.82**	-.01	.88*	.00	.97	-.05	.98	-.21	.87	-.04	.76	.07
252	Love	.72	.18	.69	.14	.69	.22	.97	-.05	.92	.12	.79	.06	.86	-.03
251	Values	.96	-.11	.86	-.02	.92	-.04	.96	-.01	1.05	.00	.90	-.05	.82	.03
206	Sex	.82	.04	.83	.01	.86	.06	.76*	.18	.98	-.06	.81	.04	.87	-.04
280	Interests	.99	-.11	1.01	-.21	1.02	-.15	1.02	-.08	.99	-.16	.95	-.12	.62**	.26
245	Conflicts	.65**	.25	.70	.18	.74	.18	1.07	-.15	.81	.20	.67*	.21	.75	.10
182	*Emotional stability*	.70*	.22	.73	.14	.78	.16	1.07	-.13	.83	.23	.70*	.19	.90	-.10
168	*Openness*	1.03	-.12	.95	-.15	1.02	-.12	1.04	-.07	1.10	-.19	.96	-.11	.73	.12
392	*Extraversion*	.89	.00	.95	-.12	.96	-.07	1.03	-.12	.80	-.04	.80	.06	.68*	.19
238	*Agreeableness*	.63**	.29	.69	.18	.70*	.23	1.08	-.15	.80	.23	.67**	.19	.89	-.08
49	*Conscientiousness*	.94	-.06	.89	-.07	.93	-.04	.99	-.06	1.02	-.08	.93	-.11	.82	.02

Notes. *Different sets of variables were included to evaluate their relevance in predicting different measures of relationship quality:*
Items and scales of actor effects (P1)/ partner effects (P2)/ similarity effects (Sim);
Items and scales of relationship-related personality (R. pers.)/ general personality (G. pers.);
r²: forecasting coefficient of determination note that, since model training and model evaluation are carried out on different data sets, r² may become negative.
MSE: mean squared error as deviations between the values which were predicted by the model trained on the train data and the real observed values of the test data; an MSE=1 reflects a deviation of one SD because all variables were z-standardized.
Performance values in the bottom of the table had to become better to become significant due to because there were less variables the model was based on.
nb.: number of variables the models selected from.
* p<.05, **p<.01, ***p<.001 significantly better than the baseline model predicted the relationship quality measure.

Table 8: 10x10-fold cross-validation-performance of the elastic net models predicting subscales of harmony after four years (n=192)

nb.	Predictor variables	Relationship notions		Arrangement in Corporate Domains		Positive emotions		Arrangement in Future Domains		Problem solving	
		MSE	r²	MSE	r²	MSE	r²	MSE	r²	MSE	r²
	baseline	1.04	-.11	.95	-.11	1.03	-.11	.98	-.12	1.02	-.10
2665	Sim, Mod	.96	-.05	.94	-.12	.84*	.10	1.00	-.16	.96	-.06
245	Mod	1.10	-.20	.99	-.18	1.09	-.20	.99	-.14	1.08	-.20
2420	Sim	.95	-.04	.93	-.11	.79**	.14	.96	-.11	.95	-.04
5149	P1, P2, Sim, Mod	.80***	.13	.70***	.16	.68***	.26	.80**	.07	.75***	.16
4904	P1, P2, Sim	.80**	.12	.69***	.16	.68***	.27	.79**	.07	.74***	.18
2484	P1, P2	.75***	.17	.71***	.13	.64***	.30	.80*	.07	.79**	.11
1242	P1	.74***	.18	.69***	.16	.72***	.22	.82*	.05	.79***	.12
1242	P2	.94	-.03	.92	-.09	.82**	.09	.91	-.07	.86	.03
4423	R. pers., G. pers.	.80**	.12	.70***	.15	.71***	.24	.80*	.07	.74***	.17
3177	R. pers.	.78***	.15	.68**	.16	.65***	.30	.77**	.10	.67***	.24
1246	G. pers.	.98	-.07	.90	-.07	1.00	-.09	1.03	-.18	.95	-.04
252	Love	.92	.02	.75	.11	.90	.01	.93	-.14	.81	.10
251	Values	1.05	-.13	.97	-.10	.96	-.03	.98	-.10	.93	-.03
206	Sex	.98	-.06	.96	-.10	.79	.15	.97	-.11	.94	-.04
280	Interests	.99	-.06	.97	-.10	1.09	-.18	.98	-.10	1.00	-.07
245	Conflicts	.81	.11	.72*	.17	.82	.10	.84	-.01	.87	.04
182	*Emotional stability*	.83	.09	.79	.10	.88	.04	.86	.00	.87	.07
168	*Openness*	1.10	-.20	.97	-.13	1.10	-.20	1.03	-.18	1.04	-.13
392	*Extraversion*	.80	.14	.90	-.05	1.00	-.08	.96	-.10	.93	-.01
238	*Agreeableness*	.80	.14	.67*	.22	.80	.13	.81	.08	.81	.11
49	*Conscientiousness*	1.02	-.08	.97	-.12	1.00	-.09	.98	-.10	1.00	-.11

Notes. *Different sets of variables were included to evaluate their relevance in predicting different measures of relationship quality:*
Items and scales of actor effects (P1)/ partner effects (P2)/ similarity effects (Sim);
Items and scales of relationship-related personality (R. pers.)/ general personality (G. pers.);
r²: forecasting coefficient of determination note that, since model training and model evaluation are carried out on different data sets, r² may become negative.
MSE: mean squared error as deviations between the values which were predicted by the model trained on the train data and the real observed values of the test data;
an MSE=1 reflects a deviation of one SD because all variables were z-standardized.
Performance values in the bottom of the table had to become better to become significant due to because there were less variables the model was based on.
nb.: number of variables the models selected from.
* p<.05, **p<.01, ***p<.001 significantly better than the baseline model predicted the relationship quality measure.

2.3.2.1. Reproducible predictive power

As hypothesised, the cross-validated model with actor, partner and similarity variables was replicated and explained more than 25% of the variance of *RQ overall* (MSE=.55, r^2=.37, p<.001). By way of example, figure 2 shows the relation between the predicted and the actual *RQ overall* values for one of the 10 CV iterations[10].
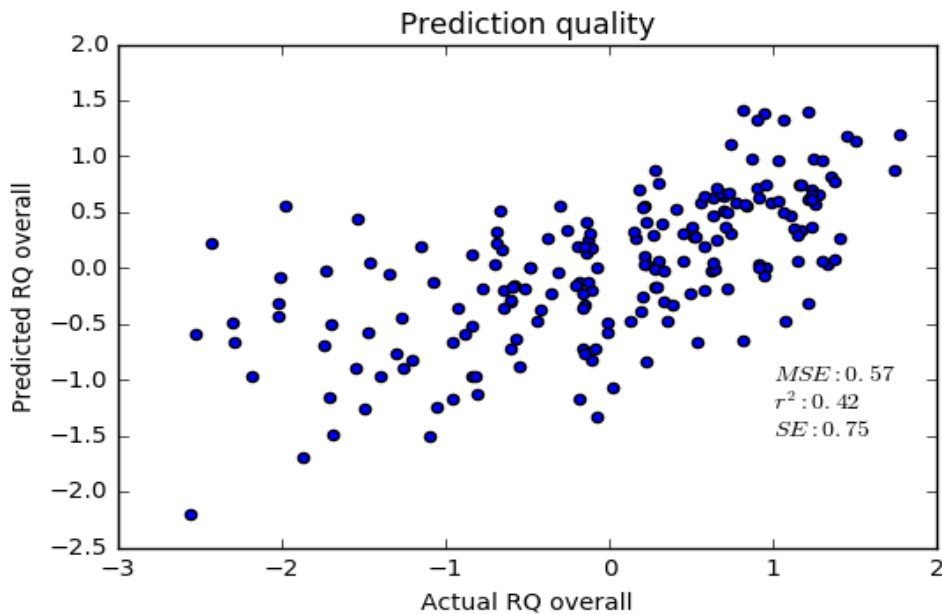
Figure 6. Actual vs. predicted RQ overall for one (!) of the 10 cross-validation iterations based on actor, partner and similarity variables.

Notes. RQ: relationship quality. MSE: mean squared error. Since only one and not the average of all 10 cross-validation iterations is presented in the figure, r^2 and MSE differ from the performance reported in Table 7.

2.3.2.2. Actor, partner and interaction effects

Neither partner nor similarity effects had a significant additional value after accounting for actor variables (Plus partner variables: t(99)=1.57, p=.119. Plus similarities variables: t(99)=.0567, p=.955). Partner variables only explained 0 to 7% of the variance for the RQ measures and predicted 5% incremental variance beyond actor effects. They had lower predictive power compared to actor variables for every RQ measure: for *RQ overall*, they significantly differed from one another (t(99)=3.78, p<.001). Contrary to the hypotheses, neither similarity variables nor additional moderators enhanced the prediction power et. al. (all three models: MSE=.55, r^2=.37).

---

[10] Appendix I.10. presents the visualizations of the results of all 10 CV iterations.

2.3.2.3. Relationship-related and general personality

(a) As hypothesised, variables of *general personality* did not have significant predictive power in addition to *relationship-related personality* variables: while the difference between models based on general vs. general plus *relationship-related personality* was significant (t(99)=5.25, p<.001), *general personality* variables had no relevant effect in addition to *relationship-related personality* (t(99)=-.553, p=.582). Overall, *general personality* had a lower predictive power for all RQ measures than *relationship-related personality* throughout this analysis: For *RQ overall*, they significantly differed in their predictive power (t(99)=5.09, p<.001).

(b) As hypothesised, models based on conflict-related (MSE=.65, r²=.25, p=.008) variables were more predictive than models based on value-related (MSE=.96, r²<.01, p=.782, n.s.) and interest-related (MSE=.99, r²<.01, p=.948, n.s.) variables. Models based on sex-related (MSE=.82, r²=.04, p=.228, n.s.) and love-related (MSE=.72, r²=.18, p=.063) attributes did not predict significantly better than the baseline.

Models based on variables of *agreeableness* (MSE=.63, r²=.29, p=.005) and *emotional stability* (MSE=.70, r²=.22, p=.042) were significantly more predictive than the baseline, while models based on variables of *conscientiousness* (MSE=.94, r²<.01, p=.702, n.s.), *extraversion* (MSE=.89, r²<.01, p=.403 n.s.) and *openness* (MSE=.1.03, r²<.01, p=.845, n.s.) were not.

The differences between the model based on conflict-related variables compared against the one based on value-related variables (t(99)=-2.30, p=.023), and the one based on interest-related variables (t(99)=-2.50, p=.014), were significant. Moreover, the model based on *openness* significantly differed from the one based on *emotional stability* (t(99)=2.22, p=.029), as well as from the one based on *agreeableness* (t(99)=3.22, p=.002). In addition, the model based on *agreeableness* significantly differed from the model based on *extraversion* (t(99)=-2.05, p=.043), as well as from the one based on *conscientiousness* (t(99)=-2.56, p=.012). Differences between the other models were not significant (t(99)|<2.0, p>.05).[11]

---

[11] Appendix I.9. presents the t-test results between the models based on different personality domains in detail.

2.3.2.4. Explorative analysis on different RQ measures

The following observations were made regarding the prediction of the different main RQ measures:

- The RQ main measures - *separation intents* (MSE=.67, r²=.16, p<.001), *relationship satisfaction* (MSE=.69, r²=.21, p<.001), *Sexual satisfaction* (MSE=.72, r²=.24, p<.001), and *harmony overall* (MSE=.60, r²=.28, p<.001) - could be predicted to a similar extent by models based on different sets of variables (see *Table 6:* upper half), while *conflicts* could not be significantly predicted better than the baseline (MSE=.88, r²=.01, p=.172).

- Descriptively, *RQ overall* could be predicted slightly more accurately than the RQ measures it was generated from.

- Facets from *emotional stability*, *agreeableness,* and conflict-related attitudes significantly predicted the main RQ measures *conflicts, separation intents, relationship satisfaction,* and *harmony overall,* but not *sexual satisfaction* (MSE>.95, r²<.01).

- Only *Sexual satisfaction* was relevantly predicted by sex-related attributes (MSE=.76, r²=.18, p=.050); when the other way around, sex-related traits did not explain significant variance of any other main RQ measure.

- *Separation intents* could be predicted significantly better than the baseline at T2, but not at T1.

The following observations were made regarding the prediction of different subscales of partnership *harmony*:

- Among *harmony* scales, *future arrangement* was the least predictable (MSE>.77, r²<.10), while *positive emotions* could be best predicted best (MSE>.64, r²<.30).

- *Future arrangement* and *problem solving* could not - or only exceedingly slightly - be predicted by models when only based on a specific trait or domain (MSE>.80, r²<.12).

- Following the pattern of the main RQ measures, *arrangement in corporate domains* was similarly high predicted by *emotional stability*, *agreeableness*, and conflict-related attitudes.

2.3.3. Variable importance

To evaluate the importance of different variables, the average weight of each variable was calculated over the 100 models (from the 10x10-fold CV) which were trained using all variables. The most important variables of each variable group (actor, partner, similarity) are

presented in *Table 5*[12]. To evaluate the importance of the different variable groups, the weights of the variables were added up. The weights of actor variables contributed 34.7% to the calculated sum, while partner variables and similarity variables account for 14.6% and 50.7%, respectively. Contradicting the result that similarity variables had no significant incremental power, they accounted for roughly half of the weights in the trained models which were based on all variables.

The most important scales mostly were those which significantly correlated with *RQ overall*. Many of the most important items belong to those scales (*Table 6* vs. Table 9). Nine of the 25 top actor variables (17.63% of all importance), five of the 15 top partner variables (3.33% of all importance) and 14 of 55 top similarity variables (12.13% of all importance) refer to previous tendencies to react positively or negatively to partnerships, e.g., the actor item "*I have every reason to be dissatisfied with my previous relationships*".

Table 9: Variable importance of top predictors from the elastic net model predicting relationship quality (n=192)

Actor variables (altogether account for 34.7%)			%	*β*
BB-PI Item	546	*I have every reason to be dissatisfied with my previous relationships.*	7.66	-.2137
BB-PI Item	302	*I often feel sexually rejected by partners.*	2.31	-.0644
BB-PI Scale	NA046	*Negativistic- prejorative*	2.15	-.0601
BB-PI Item	224	I often find exotic-looking subjects very beautiful.	1.98	.0552
BB-PI Item	537	I look for the confirmation of my sexual attractiveness from others.	1.85	-.0516
BB-PI Item	449	*Overall, I am very grateful for my happy relationship experiences thus far.*	1.68	.0468
BB-PI Item	187	*I still carry a big unfulfilled yearning for love in me.*	1.24	-.0347
BB-PI Item	524	It is important that the man has good manual skills and the woman can run a household.	1.09	.0305
BB-PI Item	631	I like to go out in the evening and want to attend a lot of events and social gatherings – with or without my partner.	1.09	-.0304
BB-PI Item	460	*Love often leaves me very confused.*	.90	-.0251
BB-PI Item	504	*I am distrustful regarding promises of partners because too many have been broken.*	.89	-.0248
BB-PI Item	240	I think that conversations with partners about relationship problems rarely lead to clarification of these.	.81	-.0227
BB-PI Item	352	I often and gladly pay partners compliments which go down well with them.	.74	.0205
BB-PI Item	653	I think, also in love the statement is true: "The more you have to offer the more you get".	.54	-.0151
BB-PI Item	76	*Traumatic experiences from my childhood still stress my partnerships today.*	.54	-.0150
BB-PI Item	135	Also in a partnership, I need a lot of time for myself.	.53	-.0149
BB-PI Scale	AG161c	Adaptiveness (capability)	.42	-.0116

---

[12] Appendix I.11 presents the whole list of the 924 selected variables with their variable importance.

Actor variables (altogether account for 34.7%)			%	β
BB-PI Item	598	I had romantic relationships with totally different personality types.	.39	-.0108
BB-PI Item	90	My partner should fit all expectations regarding order and cleanliness.	.27	.0076
BB-PI Scale	GZ049	*Happy-satisfied*	.26	.0072
mix Scale	ZU054c	Listening (capability)	.82	.0228
PD-I Item	187	*I am unhappy with myself and my life.*	.45	-.0127
PD-I Item	68	I have a demanding technical hobby.	.39	-.0109
PD-I Scale	IQ158c	Problem solving skills (capability)	.33	-.0092
PD-I Item	150	I love exactness and precise working.	.28	-.0078
Partner variables (altogether account for 14.6%)			**%**	**β**
BB-PI Item	*2*	*I often suffer from sexual dissatisfaction and imbalance in partnerships.*	1.76	-.0491
BB-PI Item	450	After an argument, I have often treated my partner as if they were not there.	1.60	-.0447
BB-PI Scale	*SU002*	*Sexually dissatisfied*	.70	-.0196
BB-PI Item	*607*	*I constantly feel the need to put every partnership to the test.*	.66	-.0184
BB-PI Item	*591*	*In my partnerships, the same patterns of relationship problems repeat themselves.*	.62	-.0172
BB-PI Item	27	I am vastly superior to other (wo)men in bed.	.51	.0142
BB-PI Scale	AO069	Demanding - optimizing	.37	-.0102
BB-PI Item	222	I am often tempted to be unfaithful.	.29	-.0082
BB-PI Scale	*KU053*	*Disturbed communicating-misunderstood*	.29	-.0080
BB-PI Item	104	I think if you love somebody you can nearly forgive everything.	.21	-.0058
BB-PI Item	217	I always feel when the partner has something on his/her mind.	.20	.0055
mix Scale	CO163p	*Conscientiousness* (preference)	1.39	.0389
PD-I Item	133	I am happy if I have no one around me.	.39	.0108
PD-I Item	70	I attach value to efficient budgeting.	.28	-.0077
PD-I Item	173	I don't give up after defeats.	.21	.0058
Similarity variables (altogether account for 50.7%)			**%**	**β**
BB-PI Item	*DIS 526*	*I am able to give myself unconditionally sexually and emotionally to a partner.*	2.63	-.0733
BB-PI Scale	*CC low-low SU002*	*Sexually dissatisfied*	2.36	.0660
BB-PI Item	*DIS 76*	*Traumatic experiences from my childhood burden my partnerships through to the present day.*	2.11	-.0590
BBP Scale	CC low-low LOVE	Sum: love-related items where both have low values	1.73	.0483
BB-PI Scale	DIS SS134_cap	Social supportiveness in a partnership (capability)	1.64	-.0457
BB-PI Scale	CC high-high EA052	Arrogant- presumptuous	1.61	.0450
BB-PI Scale	DIS DB012c	Distance in a partnership (capability)	1.51	-.0420
BB-PI Scale	CC low-high LK041	Motivated to achieve – carrier-orientated	1,16	-.0323
BB-PI Scale	CC low-low BV105	Protective - defending	1,08	-.0302
BB-PI Scale	CC high-high VE111	Understanding - sensitive	.91	.0253
BB-PI Scale	CC high-low GA038p	Emotionless - alexithymic (preference)	.87	-.0243
BB-PI Scale	*CC low-low KU053_Passive*	*Passive: disturbed communicating-misunderstood*	.82	.0229
BB-PI Scale	MOD KR160c	Free mind/fantasy/dreamer (capability)	.82	.0228
BB-PI Item	DIS 212	I find a career-based separation from the partner to be well bearable.	.72	-.0201
BB-PI Scale	CC low-high AG161cMin	Stubbornness	.58	-.0163
BB-PI Scale	CC high-high VE111c	Understanding – sensitive (capability)	.58	.0162
BB-PI Scale	MOD DB012c	Distance in a partnership (capability)	.58	.0161
BB-PI Scale	*CC low-low MB004*	*Suspicious - insecurely attached*	.57	.0160

Similarity variables (altogether account for 50.7%)			%	β
BB-PI Item	*DIS 623*	*I look for luck in relationships, do a lot for it and find it than.*	.54	-.0150
BB-PI Item	DIS 425	I repeatedly had problems with partners because I go out to less and rather like to hang around limply.	.52	-.0146
BB-PI Item	DIS 109	I find unbearable if a partner has an emancipation-hostile attitude.	.52	.0146
BB-PI Scale	*CC low-high DH073*	*Dependent - helpless*	.49	-.0137
BB-PI Item	DIS 450	I often treated an own partner as if they were not there.	.46	-.013
BB-PI Scale	*CC low-high SU002*	*Sexually unsatisfied*	.46	-.0127
BB-PI Scale	CC low-high LK041a	Attitude: motivated to achieve – carrier-orientated- lack of time for partner	.44	-.0124
BB-PI Scale	CC low-high BH086	Relationship questioning	.41	-.0116
BB-PI Scale	*MOD DH073*	*Dependent - helpless*	.37	-.0103
BB-PI Item	*DIS 591*	*In my partnerships, the same pattern of relationship problems repeats.*	.33	-.0092
BB-PI Item	DIS 521	In a partnership, I wish a lot of cultural flair.	.31	-.0086
BB-PI Scale	MOD US035	Independent - self-determined	.30	.0085
BB-PI Item	DIS 373	I am happy if a partner takes difficult decisions for me.	.30	.0085
BB-PI Scale	CC high-high BB093	Preserving relationships	.30	-.0083
BB-PI Scale	*MOD UG061*	*Weary - bored*	.30	.0082
BB-PI Scale	*CC low-high BU042*	*Unwilling to bond oneself*	.28	.0079
BB-PI Item	DIS 246	In times of need, I would do anything to get the partner and the family through.	.28	-.0077
BB-PI Item	DIS 216	I tend to curse loudly and heavily when in conversations with partners.	.26	.0072
BB-PI Item	DIS 222	I am often temped to become unfaithful.	.23	-.0064
BB-PI	DIS 667	I find sex with technical tools is disgusting.	.22	.006
BB-PI Scale	CC low-low AG161c-	Stubbornness (incapability)	.21	.0059
BB-PI Item	*DIS 84*	*Fear of failure regarding physical love making can become so strong with me that nothing works out any-more.*	.21	.0058
BB-PI Scale	CC low-high LF097	Fixed on lover's qualities	.20	-.0056
mix Scale	DIS TR141	Faithfulness (attitude)	1.98	-.0553
mix Scale	MOD TR141	Faithfulness (attitude)	1.94	.0541
mix Scale	DIS KI013i	Children and family (attitude)	1,20	.0335
mix Scale	*CC low-high EI159min*	*Anxiety*	.66	-.0184
mix Scale	CC low-low RE143_1	Traveling	.46	.0129
mix Scale	MOD LQ200p	Lover's qualities (preference)	.42	-.0117
mix Scale	MOD EF138p	Erotic fantasy (preference)	.39	.0108
mix Scale	CC low-high AG161c	Forgiving (capability)	.35	-.0098
mix Scale	DIS RE143	Traveling	.26	-.0072
mix Scale	CC low-high DO164p	Partner dominance (preference)	.22	-.0061
PD-I Item	DIS 208	I tend to experience a fright stage and anticipatory anxi-ety.	1.34	-.0375
PD-I Item	DIS 250	I would never dare to do extreme sports.	.31	-.0088
PD-I Item	DIS 195	I'm unkind to people I find disagreeable.	.28	-.0078
PD-I Scale	CC low-high AS079p	Adventurous - sporty, experience seeking (preference)	.27	-.0075

Notes. *Table contains variables explaining >.2% of the relationship quality overall on average. Overall model: mean squared error=.55, r²=.37. 925 of 4904 variables selected. β: standardized regression beta.*
CC: Combination Counts for all items of one scale. DIS: distance between partners values. MOD: Moderator of partner's z-values of scales. Variables referring to previous tendencies to react emotionally positive or negative to partnerships are written in cursive type. PD-I: Personality Domain Inventory. BB-PI: Attachment- and Relationship-related Personality Inventory. mix: scales with mixed items from PD-I and BB-PI.

2.4. Discussion

2.4.1. Conclusion

2.4.1.1. Reproducible predictive power

The ML approach added to the general power and reproducibility of predicting RQ with personality data longitudinally: 37% of the *RQ overall* measure could be explained after four years in CV. Compared to former studies using simpler correlative analyses with personality data (Dyrenforth, Kashy, Donnellan, & Lucas, 2010; Malouff, Thorsteinsson, Schutte, Bhullar, & Rooke, 2010), this is a significant improvement. Follow-up studies could now examine whether the prediction works similarly well for future rather than current relationships only, especially for *break-up* as a dichotomous outcome.

2.4.1.2. Actor, partner, and interaction effects

Actor effects alone explained nearly all variance of the RQ measures, while partner or interaction variables did not have an additional effect; this way, the results of the current application of ML correspond with those found by traditional regression approaches (Dyrenforth, Kashy, Donnellan, & Lucas, 2010). While actor and partner effects explained similar variance (18% compared to 27%) when predicting romantic attraction with the help of ML in a small previous study (Joel, Eastwick, & Finkel, 2017), actor effects were more predictive for later RQ in the current one (33% compared to 7%): initially being attracted to somebody attractive might more correspond more with their characteristics than becoming happy with them later; but both initial attraction, as well as later RQ, might be influences by one's own traits to a similar extent.

Partner effects incrementally explained 4% of RQ's variance beyond actor effects, which corresponds with effect sizes of former larger longitudinal studies (Becker, 2012; Dyrenforth, Kashy, Donnellan, & Lucas, 2010), while similarity effects on top of actor and partner effects added no predictive value. The different methods used to scale similarity (distances, moderators, CCs) were also not able to additionally contribute. Possible explanations are:

(1) Interaction effects do not truly have a relevant incremental predictive value beyond actor effects, or they are so minor that they could not be detected with the current small sample. Nonetheless, possible reasons why the similarities are correlated with RQ may be due to their correlations with relevant actor and partner effects.

(2) It could be that similar partners evolve more functional coping strategies with one another, or that a functional personality is more likely to look for similar partners; this way, the actor effects would produce similarity effects, which then explain the same variance of RQ.

(3) The BB-PI and the PD-I might not be appropriate tests for measuring relevant interaction effects.

(4) Since interaction effects only combine information that actor and partner variables already provided, they might not have additional predictive power, when they are added. In this case, the linear regression working with the different scores could be an inappropriate method to simulate a potential interplay between actor and partner effects on RQ.

The high variable importance of similarities could result from the fact that interaction effects contain more (relevant) information than separate actor and partner effects. The models applied both lasso and ridge penalties, which - as a duo - are trained to select the most informative predictors among highly correlated ones. Therefore, interaction variables might have been more likely chosen.

2.4.1.3. Relationship-related and general personality

Replicating former results (Noftle & Shaver, 2006), models based on *general personality traits* predicted RQ less effectively than models based on *relationship-related personality traits*. Furthermore, *general personality* had no additional significant predictive power when taking *relationship-related personality* into account. General *personality traits* might only significantly influence the quality of a partnership when they directly affect interpersonal coping, e.g., are attached to social skills or are experienced in such relationship surroundings as it is the case for *agreeableness* or neuroticism; both are directly linked to interpersonal conflict coping. While neuroticism includes the tendency to experience negative emotions during conflict, *agreeableness* contains a set of functional and dysfunctional coping strategies for interpersonal issues and situations. Correspondingly, non-conflict-related attitudes such as general values and interests, *openness*, and *conscientiousness* do not seem to play a significant role for RQ et. al.. Even *extraversion*, which refers to interpersonal contact but not to interpersonal conflict, does not play a major role for RQ.

This way, the present work succeeded in replicating results found with data from behavioral observations (Gottman J. M., 2014) with data from self-assessment: particularly, communication and conflict-related *personality traits* predict *break-up* and relationship hap-

piness, but not *sexual satisfaction*. The present work indicates that these characteristics might at least partly be consistent across different relationships. This idea is supported by the finding that questions about the quality of former relationships were among the most important predictors. This general competency in relationship is represented within the love-related and conflict-related variables that reveal to be important for nearly every part of RQ.

2.4.1.4. Explorative analysis on different RQ measures

RQ overall, relationship-satisfaction, *separation intents*, *harmony overall*, positive emotions, problem solving, arrangement in basic, and corporate domains could all be similarly well predicted and their model outcomes followed a very similar pattern, thus indicating that these reflect the more general and overlapping aspects of the same construct. Deviant patterns occurred for:

- *Conflicts*

  In contrast, models based on different personality domains (A, N, conflict-based traits) failed to predict perceived interpersonal conflicts. *Conflicts* could not even be significantly predicted by actor effects, but – surprisingly - by partner effects only. It is possible that *conflicts* caused by one party are not seen as such by that party; this could be an interesting topic for future research.

- *Sexual satisfaction*

  Models based on sex-related variables specifically succeeded in predicting *sexual satisfaction* – an effect which does not spread to other measures of RQ. Reversely, models based on other personality domains did not significantly predicted *sexual satisfaction*. The exclusiveness of these predictive successes separates *sexual satisfaction* from the main components of RQ.

- **T1** *separation intents*

  The design - cross-sectional vs. longitudinal – was shown to be a significant moderator for the sizes and directions of predictions concerning RQ for the *Big Five* (Malouff, Thorsteinsson, Schutte, Bhullar, & Rooke, 2010). Indeed, the present results also indicated strong differences between model performances of T1 and T2 models.

2.4.1.5. Predictive validity of combination Counts

The predictive validities of CCs indicated that different combinations of partner's values affect RQ differently for various traits. Three patterns could be observed and require further exploration: (1) shared strengths often affected RQ positively, while shared weaknesses

often affected them negatively; (2) certain differences were only negative when one partner scored high and the other scored low, which could be due to preference non-fittings when one partner prefers a trait that the other does not have.

2.4.2. Limitations and outlook

In the following sections and in *Table 6*, the limitations and benefits of the present study are juxtaposed and discussed. In summary, future work should contribute to further improvements in predictions of RQ and to increased generalizability in the models developed.

Table 6: Study I evaluation

	Benefits	Limitations
Generalizability	+ Longitudinal design enabled prediction over time. + Immanent *cross-validation* of models protected from overfitting.	- The sample size was restricted. - Only German couples were examined. - Only partnerships already existed at Time 1 were assessed.
Model fit	+ The *elastic net* with optimization coefficients alpha and lambda could cope with large number of highly correlated variables. + Dyadic data was handled by allocating partners of one dyad to either train or test sample.	- The large number of variables in proportion to the sample size restricted the model fit. - Only linear effects were analyzed.
Comparability	+ Models for variable sets and outcomes were systematically juxtaposed.	- The number of variables the models selected from and the number they selected varied.

2.3.2.1. Sample

Immanent CV of models protects from overestimating the predictive power and enhances replicability. Nonetheless, the exclusively German sample restricts generalizability of the results across different cultures. The relatively small sample size – additionally including 7.0% cross-sectional data within the longitudinal sample (27 partners who attended at T2 but not at T1) – could also have limited predictive power, especially due to the comparably high number of variables. Additionally, since the couples existed at T1, partners of the current sample already influenced one another, e.g., might have changed their partner preferences or their self-perception, based on their relationship with the current romantic partner. This might affect the applicability of the models for partner matching on singles. Although general and *relationship-related personality traits* turned out to be more robust over time than relationships are (Kirkpatrick & Hazan, 1994), it could still be the adaptable, non-stable variance in these trait measurements which are correlated with RQ, e.g., the assessment of the own competency in relationships rises with quality of the current relationship (see *Study II: Relation-*

ship-related Long-term-stability of Relationship-related Personality Traits). To fully ensure applicability in, e.g., the mating context, future work has to replicate the models in samples of potential partners who get to know each other after they take the personality test.

2.3.2.2. Study design

Although the current longitudinal design enables prediction over a four-year term at least, longer-term examinations would still be desirable. An advantage of the present study in terms of comparability is the systematic juxtaposition of models with different variable sets and outcomes. Still, the number of variables, the models selected from, and the number finally selected varied, making a direct comparison between the models difficult: Prediction typically increases in stability with higher numbers of predictors and is thereby more easily significant in comparisons.

Some preceding studies indicated that shared method variance in dyadic data analysis can lead to differences in prediction quality. This has been discussed as being an especially relevant question, when the present results would be considered for a partner matching (Orth, 2013). The algorithms in the present study solve this issue by allocating the partners of the same dyad both either to the train sample or both to the test sample for every iteration of the CVs.

Through the interdisciplinary synthesis of data science methods and psychology, the present study managed to contribute with cross-validated, accurate and longitudinal predictions. The *elastic net* managed to cope very well with the large amount of highly correlated variables due to the different optimization coefficients, alpha and lambda. Future studies could examine the possibility of unexplained non-linear personality-RQ association, such as those studied by Hudson & Fraley (2014) or Joel, Eastwick and Finkel (2017) through the application of non-linear ML methods such as decision-trees[13]. Moreover, models integrating aspects of the context could be interesting to further explore the situation-person interaction with the help of ML.

---

[13] Appendix I.13 presents some pilot results about decision-trees compared to linear models, both predicting *break-up* and *RQ overall* by BB-PI items, scales and partner differences.

3. **Study II***: Relationship-related Long-term-stability of Relationship-related Personality Traits*

Abstract

To evaluate how relationships and personality influence one another, the present study examined the changeability of *relationship-related personality traits* over time and over different relationship statuses. Thereto, participants completed a questionnaire for self-assessment twice, with a four year interval between the two occasions. Various facet changes over time were tested and described.

The moderate retest-reliabilities for facets and the high retest-profile-correlations suggest a certain robustness; at the same time the findings indicate that current relationship experiences significantly influence the perceived relationship-related personality: 40 of the 90 facets differed, in how (much) they have changed over time, between subjects who end an old relationship, start a new one, continue the current one or remaining without a relationship. Most facets linked to *tendencies to react emotionally positive or negative* were among these and may be confounded with the quality of current relationships. By contrast, most facets linked to sexual preferences (8 of 11) were especially robust over different relationship surroundings. Practical implications for applying questionnaires to predict *relationship quality* of potential future couples are discussed.

Keywords: dissolution, close relationships, longitudinal studies, *personality traits*, stability, *relationship satisfaction*, singles

3.1. Background

3.1.1. Stability of relationship-related personality traits over time

Whilst models for couples' communication styles only claim relationship-specific va-lidity, *relationship-related personality traits* are moderately stable over a lifetime, as are in-dividual differences in one's *general personality* (Kirkpatrick & Hazan, 1994; Roberts, Walton, & Viechtbauer, 2006). General and *relationship-related personality traits* are shown to be more robust over time than relationships and to become more robust with age (Kirkpatrick & Hazan, 1994; Roberts, Walton, & Viechtbauer, 2006). As a support for part-nership-overarching influences, negative effects of high neuroticism and low *agreeableness* on the RQ seem to be consistent across different relationships, since they predicted dissatis-faction - not only for current partners, but also for future ones (Donnellan, Larsen-Rife, & Conger, 2005; Robins, Caspi, & Moffitt, 2002). Nonetheless, the partner is a central aspect of an individual's social environment; hence it is hardly surprising that significant changes in partnership situations have been associated with significant personality changes (Caspi & Herbener, 1990). It has been suggested that a consistent environment is a main cause of per-sonality consistency (Cairns & Hood, 1983; Caspi & Roberts, 1999; Moss & Susman, 1980). Correspondingly, Caspi and Herbener (1990), e.g., found that partners who were similar to each other were, in turn, more consistent over time in their *personality traits*.

3.1.2. Retest-stability for the Relationship- and Attachment-related Per-sonality Inventory

Andresen's (2012) model of *relationship-related personality* (see *Figure 7*) claims to be generalizable over different partners, as well as over time, and to exhaustively contain the "Sexy Seven" (Schmitt & Buss, 2000), attachment styles (Bartholomew & Horowitz, 1991; Hazan & Shaver, 1987), and love styles (Lee, 1988).
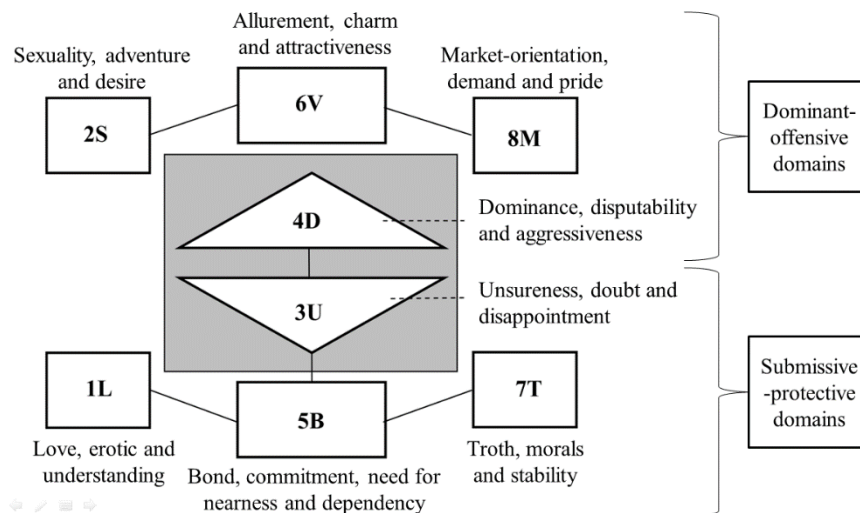
Figure 7. Eight-domain model for the *relationship-related personality* measured with the

Relationship- and Attachment--related Personality Inventory (Andresen, 2012).

Notes. *Andresen developed the eight-domain-inventory by an exploratory factor-analytic aggregation of 111 facets with the Principle Component Analyses and oblique rotation (Delta=0, standard). Only one item of one facet was used per domain, so that each domain consists of 18 items. Visualized proximity of the domains corresponds to the magnitude of inter-correlations.*

A study by Wunderlich (2011) on the BB-PI showed high retest-correlations for 142 subjects who had the same partner after a one-year period; details are presented in *Table 10*. It was assumed that the more 'you'-centered domains with rather submissive-protective contents (r were more stable than the more 'me'-centered ones with rather dominant-offensive contents, since they more attuned to adapt to the needs of a current partner or partnership. Up to now, it is an open research question, how traits might be affected in longer periods of time and how they do this for individuals who change their relationship (status).

Table 10: One-year-retest-stabilities for the eight 18-item-domains of the Relationship- and Attachment-related Personality Inventory which is applied in the current Study II (n=142)

	Item example	T1-T2 r
6V	I approach others to get confirmation that I am very sexy time and again.	.87
4D	Mostly, I argue loudly and vehemently with my partner.	.84
2S	Sexuality plays a central role in my relationships.	.84
3U	From the beginning, every potential partnership is charged with some distrust for me.	.83
8M	I would like to be proud of my partner.	.82
1L	Regarding physical love, I enjoy a long, gentle foreplay.	.78
5B	In general, I am a person who is very much needs closeness and who has great fear of *break-ups*.	.78
7T	I attach great importance to a lasting and reliable partnership.	.78

Notes. *T1-T2 r: Retest-Pearson-correlation between Time 1 and Time 2 (Wunderlich, 2011).* *For domain titles see Figure 7.*

3.1.3. Associations of partnerships to mean-level-changes in personality

Divorce, *break-up*, and marriage had no or only exceedingly small effects on stability of the *Big Five* in studies using large samples (Specht, Schmuckle, & Egloff, 2011), but particularly aspects of *relationship-related personality* seem malleable in response to changing one's partner or relationship status (i.e., entering a partnership or going back to being single). (1) In a prospective study by Kirkpatrick and Hazan (1994), experiencing a *break-up* increased the likelihood of moving from a secure to an avoidant attachment style. Conversely, experiencing more satisfying relationships was in some cases associated with a transition from an avoidant to a secure attachment style, while securely when compared to insecurely attached respondents less likely reported one or more *break-ups* during a four-year interval; this suggests that the current relationship status can be predicted from the previously reported attachment styles. To conclude, on the one hand, attachment styles work as trait-like characteristics in adults that influence the course and outcome of a relationship. On the other hand, they emerge as descriptions of the quality of one's current relationship (Caspi & Herbener, 1990), seemingly influenced by an inner working model that is evolved from current and former relationships (Kirkpatrick & Hazan, 1994; Scharfe & Bartholomew, 1994). However, a study by Scharfe and Bartholomew (1994) did not observe elevated levels of changes after relationship transitions, i.e. after *break-ups* or marriages.

Corresponding to these inconsistent results, there is still the open research question concerning how precisely changes in partnerships interfere with which (other) *relationship-related personality traits*.

3.1.4. The present study

The present study employed a four-year longitudinal design to examine the generalizability of self-assessed *relationship-related personality traits* over varying relationship surroundings and partners. Patterns for mean-level-changes and the retest-stability over time were assessed, extending the former findings retest-stability of the BB-PI-measures. The following hypotheses for the changeability of facets were tested:

(1) Correlates of individual retest-stability

(1a) Replicating existing research results, personal retest-profile-correlations will enhance with age but will not differ per gender. (1b) It will be associated with relationship experience - before, between and during the measurements.

(1c) Therefore, individuals who live in different relationship surroundings (i.e., who stay without a partnership, who begin a new one, end an old one or stay within the same one) will differ in their retest-stability.

(2) Differences between T2-T1 facet values

(2a) Some facets will decrease or increase over time.

(2b) These changes will differ between individuals who live in different relationship surroundings.

(3) Distances between T2-T1 facet values

The extent to which some facets change over time will differ between individuals who live in different relationship surroundings.

3.2. Methods

Relationship-related personality traits and demographics were measured twice within the same (dependent) sample: at T1 and at T2 with four years in between.

3.2.1. Sample

The sample consisted of 283 German adults. At T2, 144 participants (50.9%) were male and 139 (49.1%) were female. The age was m=41.3 with a SD=12.2 years (Range: 18 to 72). 150 participants (53.0%) had a university degree, 85 (30.0%) had a high-school diploma (German: Abitur), and 48 (17.0%) had finished secondary education. As visualized by *figure 2,* the medial number of past relationships was m=6.69, with SD=4.86 (Range: 0 to 40).

The current prospective study examined subjects who experienced differing changes in close relationships during the four years.

Table 11 shows the distribution of relationship statuses, partner changes across T1 and T2 and the dysjunct categorization of subsamples which will be used for the further explorations (*left column*): 64 subjects left an old relationship (22.6%), 40 started a new relationship (14.1%), 102 remained in their old relationship (36.0%) and 74 remained without a relationship (26.1%). Three subjects restarted an old partnership and thereby were treated as dropout (1.06%).

Table 11: Subsamples with different changes in relationship surroundings

Subsamples		N = 283	Relationship status Time 1 (2008)	New partner in between	Relationship status Time 2 (2012)
still-singles n=74	lonely single	41	single	no	single
	again-singles	33	single	yes	single
new-partners	*new-partners*	40	single	yes	committed
	dropout: back to ex	3	single	no	committed
still-partners	still-partners	102	committed	no	committed
break-ups n=64	lonely *break-ups*	18	committed	no	single
	partner-changers	25	committed	yes	committed
	break-ups again	21	committed	yes	single

Figure 8 and *Figure 9* show the distributions for number of relationships before the first measurement and between the measurements: More than half of the participants did not start a new relationship between T1 and T2; most participants had between two and eight former partnerships.
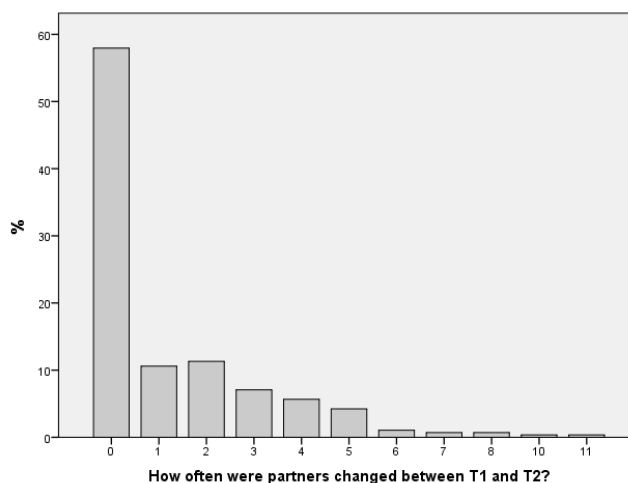
Figure 8. Histogram for the number of partnerships between the two measurements (n=280).
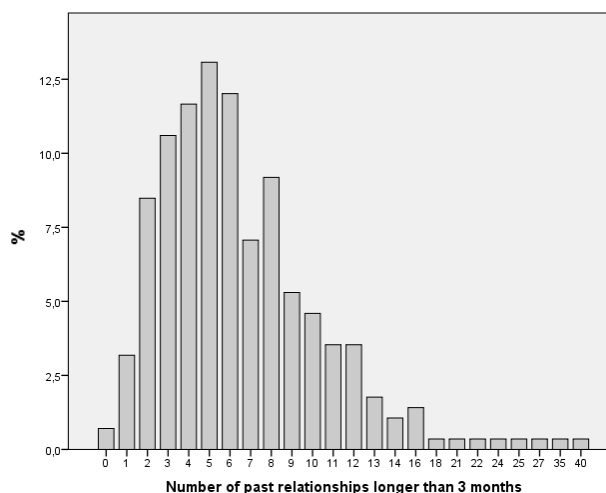
Figure 9. Histogram for the number of partnerships before the first measurement (n=280).

3.2.2. Operationalization

Relationship-related personality characteristics were measured with the BB-PI (Andresen, 2012), as described in *Study I* (*see Time one personality*), but for the present study, only the original facets that contain the sum of five items were used for analyses. 21 of the 111 facets were dropped to reduce the length of the immensely large survey and, resulting, to reduce the dropout rate. The 21 facets were deselected by rationally assessing their irrelevance for partner matching purposes[14]. For instance, the facet *culinary delighted* (KS003) was excluded. Cronbach's Alpha for the remaining 90 facets ranged from .67 to .91.

3.2.3. Procedure

Per person, an individual stability parameter was calculated as a correlation between T1 and T2 values over the 90 facets; this individual retest-stability was correlated with demographic characteristics. To isolate the effects of the relationship surroundings, the sample was divided into four disjunct subsamples (see *Table 11*). Pearson correlations between BB-PI-facets of T1 with those of T2 were composed separately for the entire sample and for the subsamples.

ANOVAS were performed:

- To detect overall stability differences between different relationship surroundings, an ANOVA with individual retest-stability as a dependent variable and subsample as independent variable, plus post hoc tests, were run.

- To examine the shifts between T2 and T2 values for all 90 facets - general and across different relationship surroundings - a mixed-design MANOVA with time as the inner-subject factor, subsamples as a between-subject-factor and age as a co-variable was conducted. The dependent variables were the 90 T2-T1 facet differences.

- To examine the general answer deviations between T2 and T1 values for all 90 facets 90 *T2-T1 facet distances* across different relationship surroundings, an ANOVA with subsample as independent variable and the T2-T1 facet-distances in facet values as dependent variables was conducted.

- Post hoc tests were run to examine which facets were robust across different relationships and which significantly changed more or less
    - when a new relationship was entered (comparing *still-singles* with *new-partners*);

---

[14] Appendix II.2. presents all facets including the non-selected ones.

o   when an old relationship was over (comparing *still-partners* with *break-ups*);

o   with different relationship statuses (comparing *still-partners* with *still-singles*).

Main effects were stated as significant if p<0.01: this strict level of significance was evaluated as appropriate due to the vast number of conducted tests. Post hoc tests were run with the method of *Least Significant Differences* and were stated as significant if p<0.05. Despite some scores were probably not being normally distributed, instead of the corresponding non-parametric methods, only ANOVAS were used. Generally, the results of these methods rarely differ and the MANOVA is the stricter treatment anyway; i.e., that chances to find larger, practically relevant effects were not reduced by this decision.

3.3.   Results

3.3.1. Preliminary analysis

Table 12 presents statistics for facet changes over the four years per different relationship surrounding, which was represented in the dysjunct subsamples[15]. These work as a preliminary data description and basis for the following MANOVAS in the later sections. As an example how to read the table, values for the facet *sexually unsatisfied* (SU002) are explained: Both measurements of SU002 (T1 with T2) correlated at r(278)=.50 for the whole sample of 280 subjects and at r(72)=.72 for the subsample of 74 *still-singles;* the facet values decreased over the four years - with an average of m=-.12 with a SD=.89 for the whole sample of 290 subjects.

---

[15] Appendix II.3. presents averages, SDs and confidence intervals for concrete scale values, *T2-T1- facet differences* and *distances* in detail.

Table 12: Comparison of mean-level changes and retest-correlations for 90 facets between different subsamples

Subsample		*Still-singles* n=74			*New-partners* n=40			*Still-partners* n=102			*Break-ups* n=64		
		m	SD	r	m	SD	r	m	SD	r	m	SD	r
NV001	In need for closeness and affiliation	-.40	.71	.59	-.53	.66	.53	-.19	1.43	.46	-.11	1.02	.18
SU002	Sexually unsatisfied	-.25	.64	.72	-.28	.92	.35	.05	.99	.49	-.16	.94	.35
MB004	Suspicious - insecurely attached	-.48	1.33	.73	-.93	1.21	.53	-.23	1.01	.67	.09	1.10	.48
UA005	Involuntarily adapted	-.18	.63	.67	-.41	.78	.53	.15	.75	.51	.18	.87	.34
SA007	Beauty conscious-attractive	-.25	.56	.78	-.34	.71	.60	-.30	.77	.52	-.48	.82	.53
IV008	Idealizing - deifying	-.17	.69	.75	-.25	1.02	.35	-.25	1.36	.35	-.09	.92	.49
ER009	Emancipatory - radical	-.10	.92	.65	-.16	.79	.34	-.43	1.46	.52	-.29	1.14	.15
FW010	Amicable - appreciative	-.66	1.72	.39	-1.05	1.78	.47	-.83	1.79	.40	-.30	1.72	.28
KA011	Physically abusive	-.2	.65	.69	-.20	.49	.75	.01	.66	.60	-.06	.90	.43
DB012	Preferring distance	-.05	.82	.53	-.04	.79	.54	-.26	1.16	.62	-.19	1.22	.28
DS014	Being a rough sexist	.01	.54	.84	-.01	.87	.46	-.15	.79	.54	-.11	.76	.73
TG015	Loyally attached	-.14	.94	.68	.07	.52	.70	-.90	1.79	.54	-.15	1.32	.32
VA016	Verbally aggressive	-.04	.73	.74	-.16	.68	.53	-1.65	.82	-	-.41	1.27	.53
ES018	Esoteric - spiritual	-.02	.67	.80	-.24	.63	.73	-.15	1.1	.56	-.30	.95	.51
PT019	Primitive - instinctively	.09	.65	.52	.00	.92	.34	-.56	1.18	.60	-.02	.71	.62
EK020	Jealous - controlling	-.22	.81	.80	-.40	.64	.70	-.30	.78	.69	-.09	.91	.51
MK021	Preferring artistic - cultural	-.07	.67	.61	-.60	.84	.59	-.14	.89	.49	-.13	.86	.48
PU022	Polygamous - unfaithful	-.02	.48	.78	-.14	.78	.48	-.05	.67	.73	-.16	.90	.62
ST023	Sexually driving force	-.11	.69	.51	.13	.94	.21	-.40	1.35	.48	-.14	.89	.50
TR024	Traditional - fixed on roles	-.01	.48	.72	-.06	.6	.54	.01	.66	.61	-.05	.91	.28
PB025	Phlegmatic - comfortable	-.03	.62	.68	.05	.79	.22	-.02	.71	.59	-.05	.70	.54
VG026	Trustful - salvaged	.01	.62	.62	.14	.76	.38	-.05	.70	.45	.03	.85	.35
MP027	Macho - boastful	-.17	.62	.58	-.01	.85	.43	.07	.86	.57	-.15	.81	.59
SS028	Self-damaging	-.39	.75	.67	-.73	.95	.44	-.42	.74	.46	-.11	.92	.39
IA031	Intimacy-aversive	-.34	.70	.32	-.50	.68	.30	-.30	.72	.49	-.33	.68	.56
RK032	Talkative - communicative	-.06	.54	.57	-.29	.40	.69	-.18	.63	.51	-.19	.82	.31
FG034	Fearful - security searching	-.04	.62	.71	.01	.67	.55	-.01	.88	.60	.03	.99	.42
US035	Independent - self-determined	.08	.68	.56	.01	.71	.46	-.08	.75	.50	-.08	.78	.44
HG037	Histrionic - compliant	-.12	.60	.77	-.38	.84	.49	-2.11	.73	-	-.52	1.13	.57
GA038	Emotionless - alexithymic	.06	.63	.65	.01	.79	.30	.07	.68	.60	-.09	.92	.51
WS039	Bold-sexually	-.13	.64	.76	-.12	1.2	.44	-2.83	.99	-	-.18	1.15	.61
KV040	Avoiding conflicts-closed	.00	.60	.59	-.2	.75	.27	.04	.83	.30	.15	.73	.37
LK041	Motivated to achieve - carrier-orientated	.02	.73	.62	-.24	.74	.72	-.40	1.23	.64	-.25	.96	.40
HE043	Hard - egoistic	.05	.50	.80	.05	.71	.56	.05	.71	.59	-.30	.83	.62

	Subsample	*Still-singles* n=74			*New-partners* n=40			*Still-partners* n=102			*Break-ups* n=64		
		m	SD	r	m	SD	r	m	SD	r	m	SD	r
SE044	Sensible - emotional	-.2	.74	.67	-.15	.69	.67	-.13	.73	.56	-.06	.88	.49
NA046	Negativistic - pejorative	-.22	.68	.80	-.56	.93	.47	-.22	.97	.44	.54	1.19	.22
SZ047	Pugnacious - angry	-.02	.71	.60	.13	.80	.53	-.34	1.33	.43	-.23	.84	.41
GZ049	Happy - satisfied	-.44	1.17	.66	-.35	1.43	.24	-.81	1.57	.63	-.66	1.4	.24
ZS050	Ensuring future	.04	.60	.66	.05	.77	.53	-.58	1.23	.54	-.18	.96	.15
TA051	Tolerant - accepting	-.08	.73	.55	-.38	1.22	.13	-.23	.78	.42	.00	.82	.22
EA052	Arrogant - overweening	-.18	.43	.74	-.15	.50	.68	.12	.60	.39	-.31	.56	.66
KU053	Disturbed communicating - misunderstood	-.23	.69	.57	-.29	.75	.55	-.21	1.04	.57	-.11	.84	.36
FD055	Leadership quality - dominant	-.03	.51	.65	.04	.52	.70	.00	.69	.44	-.02	.70	.46
EV056	Lonely - desperate	-.35	.72	.73	-.59	.98	.37	-.95	1.58	.25	-.14	1.33	.08
SN058	Stressed - nervous	.04	.58	.75	.15	.58	.78	.16	.59	.63	-.07	.85	.59
LE060	Passionate - emotional	-.08	.59	.76	-.42	.61	.74	-.40	.91	.45	-.03	.85	.45
UG061	Weary - bored	-.01	.78	.65	-.42	.98	.24	-.07	.71	.48	-.13	.99	.44
FH062	Caring - helpful	-.14	.75	.69	-.04	.58	.43	-.46	1.34	.35	-.06	1.04	.43
HD063	Domineering - dominant	.00	.50	.74	.16	.63	.62	.14	.76	.48	-.22	.78	.43
RE064	Erotic - imaginative	-.22	.82	.36	-.47	.73	.42	-.05	.73	.49	-.02	.86	.47
SG065	Shy - inhibited	-.08	.60	.75	-.18	.80	.39	-.05	.78	.54	.14	.87	.58
WA066	Workaholic	.05	.55	.68	-.02	.76	.52	.14	.61	.65	.07	.87	.47
RN068	Vindictive - vile	-.12	.77	.76	-.12	.78	.63	-.07	1.01	.46	.06	1.02	.50
AO069	Demanding - optimizing	-.03	.65	.71	-.28	.81	.54	-.81	1.18	.36	-.26	.98	.40
WI070	Daring - imposing	-.02	.61	.64	-.12	.66	.64	-.10	.73	.40	-.03	.79	.44
LB072	Qualified as lover	-.05	.52	.66	-.01	.65	.36	.18	.64	.63	-.06	.74	.47
DH073	Dependent - helpless	-.07	.46	.72	-.1	.61	.66	-.09	.66	.42	-.11	.78	.33
PA074	Punitive - antideviant	-.06	.74	.59	-.06	.88	.51	-.62	1.36	.43	-.19	.95	.27
UK075	Enterprising - outgoing	-.12	.61	.69	-.55	.73	.60	-.26	.61	.72	-.16	.95	.43
SR078	Self-assured - robust	.06	.54	.72	.07	.64	.64	.06	.53	.68	-.06	.85	.35
AS079	Adventurous - sporty	.03	.67	.68	-.34	.59	.76	-.14	.78	.56	.05	.81	.47
KW080	Church-traditional valuing	.06	.71	.69	.04	.86	.60	-.07	1.02	.44	-.08	.98	.41
DN082	Depressive - negativistic	-.02	.83	.63	-.54	.99	.26	-.28	.92	.53	.31	.96	.45
PS083	Promiscuous - partner changing	.09	.57	.82	.13	.79	.65	.00	.73	.55	.01	.71	.69
SV084	Fearing sexual failure	-.20	.86	.74	-.24	.74	.51	-.18	.89	.71	-.13	1.01	.53
BH086	Relationship questioning	-.15	.63	.78	-.29	.79	.60	-.06	.67	.68	-.34	.92	.51
MO088	Market value orientated	.31	.90	.25	.23	.90	.33	-.80	1.52	.39	.14	.88	.40
SF089	Eager to look good	-.02	.63	.62	.02	.67	.58	-.07	.73	.56	-.16	.85	.47
GE090	Demanding *conscientiousness*	.00	.46	.77	.02	.51	.71	.10	.53	.53	-.02	.59	.29
LA092	Demanding lover's qualities	-.12	.64	.56	-.11	.65	.56	-.10	.72	.55	-.13	.85	.44

Subsample	Still-singles n=74			New-partners n=40			Still-partners n=102			Break-ups n=64		
	m	SD	r	m	SD	r	m	SD	r	m	SD	r
BB093 Preserving relationships	-.27	.75	.69	-.45	.86	.63	-.05	.70	.66	-.08	1.08	.25
RL094 Romantic-love illusionary	-.16	.75	.67	-.37	.98	.45	-.17	.89	.61	-.30	1.04	.35
OF095 Obsessive - partner fixed	-.31	.67	.70	-.53	.87	.34	-.60	.93	.43	-.07	1.05	.33
SH096 Beauty-loving - hyper aesthetic	-.04	.59	.73	-.18	.74	.46	-.13	.72	.63	-.27	.86	.46
LF097 Fixed on lover's qualities	-.09	.65	.62	-.08	.70	.65	-.11	.75	.56	-.06	.88	.48
PV098 Varying in partner type	.06	.96	.73	.24	10	.41	-.39	1.25	.57	-.02	1.28	.20
PD099 Perverse - deviant	-.06	.65	.69	.15	.81	.52	-.22	.98	.45	.15	.94	.48
VV101 Willing to reconcile - forgiving	-.13	.71	.67	-.32	1.07	.44	-.04	.77	.48	-.18	.74	.34
CG102 Charming - winning	.00	.58	.72	.01	.64	.74	-.04	.99	.45	-.18	.83	.55
GH103 Good-natured – humorous - agreeable	-.02	.54	.71	-.16	.57	.59	-.15	.53	.69	.10	.67	.58
BV105 Protective - defending	-.04	.77	.63	.02	.65	.61	-.03	.77	.51	-.15	.96	.27
GF106 Habit fixed	.07	.60	.67	.30	.72	.45	-.36	1.21	.56	.13	.85	.35
UF108 Unconventional - liberal	-.04	.59	.72	-.33	.61	.62	-.28	.75	.88	-.28	.87	.62
BE109 Insulted - huffy	-.21	.67	.66	-.11	.67	.56	-.15	.79	.45	-.17	.88	.55
MG110 Materially generous	-.05	.58	.73	-.28	.71	.72	-.19	.78	.52	-.18	.82	.52
VE111 Understanding - sensitive	-.16	.47	.71	-.03	.44	.51	-.29	.67	.29	.03	.58	.48
IE112 Intercultural - exotic loving	-.08	.87	.76	.06	.97	.66	-.23	1.04	.70	-.04	1.02	.48
BS113 Possessive - proud	.06	.58	.76	-.28	.77	.60	.07	1.15	.41	-.02	.75	.56
VS114 Providing - securing	.11	.67	.56	-.07	.68	.42	-.23	1.12	.56	-.08	.81	.32
AE115 Demanding attention	-.10	.37	.69	-.12	.65	.28	-.10	.48	.64	-.01	.57	.36
ZE117 Tender - erotic	-.07	.64	.45	-.28	.80	.37	-.72	1.37	.45	-.10	.86	.35

Notes. Descriptive statistics for changes over the 4 years per different relationship surrounding.
r: retest-correlation (Pearson) between Time 1 and Time 2 measurements with four years in between.

3.3.2. Correlates of individual retest-stability

For most individuals, the average retest-profile-stability over all the 90 facets was high (T1-T2 Pearson profile correlation r(88): m=.714, SD=.188). Table 13 shows the correlations between demographics and individual retest-stabilities.

Table 13: Pearson correlations of individual retest-stabilities with demographics

	r	n	p	*Relationship-related personality*:
Age	.16**	280	.007	▪ became more stable with age,
Gender	.10	280	.103	▪ was not more or less stable for men or women,
Relationship status Time 1	-.28**	280	<.001	▪ was more stable for singles than for partners,
Relationship status Time 2	-.16**	280	.009	
Relationship status change	-.08	280	.168	▪ was not more or less stable for individuals
Partner change in between	-.05	280	.446	with different levels of relationship experience
Number of partners in between	.03	176	.738	before or in between measurements.
Amount of previous relationships	.02	280	.701	
Duration of relationship in months	.03	152	.737	▪ was not more or less stable depending on
Duration of single status in months	.06	128	.475	duration of a relationship status.

*Notes. **p<.01.*

For individual retest-stability, the between subject ANOVA revealed main effects of the subsample (F(3, 275)=11.2, p <.001, ηp^2=.109), and the age (F(1, 275)=8.24, p=.004, ηp^2=.029).Figure *10* pictures the direction and sizes of the subsample differences: the post hoc analyses indicated that the retest-stability was highest for *still-singles* (vs. *new-partners*: p=.035; vs. *still-partners / break-ups*: p<.001) and lowest for *break-ups* (vs. *new-partners*: p=.009; vs. *still-partners*: p=.039). *Still-partners* and *new-partners* did not significantly differ in their retest-stability (p=.280).
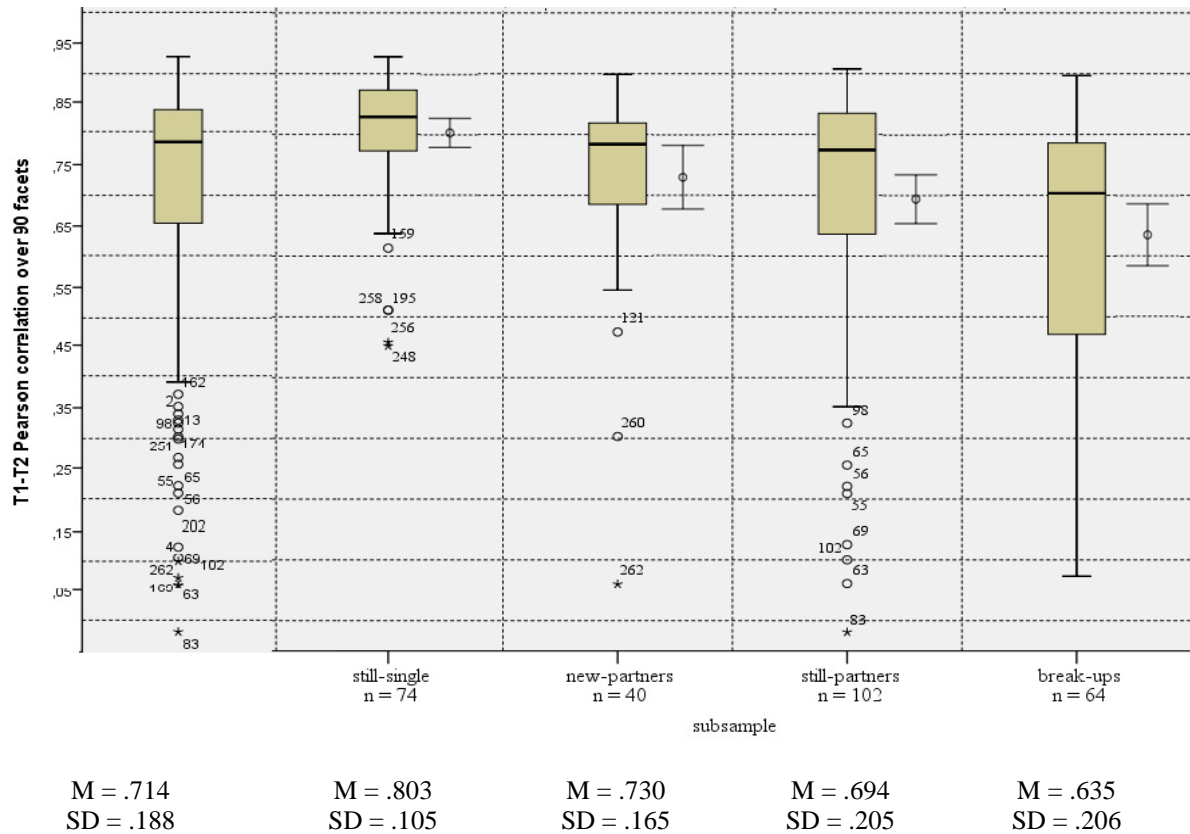
Figure 10. Distribution of individual retest-stabilities for the four subsamples.

Notes. *Comparison of the subjects who were affected by different relationship surroundings over the 4 years. Showing boxplots, error bars and confidence interval of 95% for the profile correlations between measurement one and two.*

3.3.3. *Overview about facet changes*

Retest-stabilities for the 90 facets were moderate overall (average: $r(278)=.54$ with SD=.08, range: .70 to .34). *Table 14* presents the results of the mixed-design ANOVAs on *T2-T1-facet differences* and *distances*. Significant main effects for subsample differences occur for both distances ($F=(561, 273)=2.54$, $p>.001$) and differences ($F=(275, 3)=1.80$, $p<.001$) over time. Subsample differences do not significantly interfere with age (age*differences: $F(275, 1)=1.31$, $p=.062$; age*distances: $F(275, 1)=1.32$, $p=.056$). Therefore, subjects with varying age were similarly affected by facet changes over time.

Table 14: Mixed-design ANOVA - main and interaction effects on the Pillai-Spur (n=280)

Dependent: 90 Time 1- Time 2 facet differences		F	df2	df1	p	ηp²	Nb.*
Between subjects	Subsample	1.80	275	3	<.001**	.465	21 (23.3%)
	Age	3.37	275	1	<.001**	.624	34 (37.8%)
Within subjects	Time	6.01	275	1	<.001**	.746	31 (34.4%)
	Time * Subsample	2.01	275	3	<.001**	.493	18 (20.0%)
	Time * Age	1.31	275	1	.062 (n.s.)	.392	-
Dependent: 90 Time 1- Time 2 facet distances							
Between subjects	Age	1.30	185	91	.071 (n.s.)	.389	-
	Subsample	2.54	561	273	<.001**	.553	28 (31.1%)

Notes. *Number of significant facets out of the 90, based on p<.01. ** p<.001. n.s.: not significant.*

Table 15 presents the facet changes of *relationship-related personality* facets over time[16]. The result of the exemplary facet *demanding optimizing* is explained: AO060 decreased significantly over time with an effect size of $\eta p^2 = .139$ (*df1=3, df2=276,* p<.01); it decreased more for *still-partners* than for *break-ups*, *still-singles*, or *new-partners* (SP < SS, NP, BU). For *still-partners*, the overall changes were larger than for the other subsamples (SP > BU, NP, SS); Moreover, AO060 was negatively correlated with age, i.e., the older somebody was, the lower was his / her score for AO060.

The post hoc analyses resulted in significant subsample differences: T2-T1 facet changes were besides one exception always lower for *still-singles*.

---

[16] Appendix II.4. presents main and post hoc results of between subject and inner subject MANOVAS in detail. General differences in scale values between different subsamples are also presented there.

Table 15: Sizes, directions and significance of facet changes over time (n=280)

Abbr.	Facet	Descriptives T1-T2					Innersubject Over time			Between subject Subsample * T2-T1 differences				Between subject Subsample * T2-T1 distances			
		m	SD	r	r age		F	p	ηp²	F	p	ηp²	Post hoc	F	p	ηp²	Post hoc
AO069	Demanding - optimizing	-.41	1.01	.47	-.23	***	45	***	.139	1.45	***	.102	SP<SS.NP.BU	14	***	.132	SP>BU.NP.SS
NA046	Negativistic - pejorative	-.09	1.01	.52	-.21	***	2.3		.008	13.5	***	.128	SS.NP.SP<BU	5.69	**	.058	BU>SS
BS113	Possessive - proud	-.01	.89	.55	-.14		3.96		.014	7.59	***	.076	NP.BU<SP	7.59	***	.076	SP>SS
HE043	Hard - egoistic	-.03	.71	.63	-.17	**	.53		.002	4.14	**	.043	BU<SS.NP.SP	3.93	**	.041	BU>SS
MO088	Market value orientated	-.14	1.26	.34	-.07		2.22		.008	3.91	**	.041	SP<SS	16.8	***	.155	SS<SP>NP.BU
UK075	Enterprising - outgoing	-.24	.73	.62	-.3	***	3.93	***	.1	3.31		.035	-	6.38	***	.065	BU>SS.SP
BB093	Preserving relationships	-.17	.84	.57	-.12		11.5	***	.039	2.89		.03	-	3.9	**	.041	BU>SP
HG037	Histrionic - compliant	-.97	1.2	.64	-.28	***	2.17	***	.067	2.24		.024		98.2	***	.516	SS<BU<SP>NP
ZE117	Tender - erotic	-.34	1.06	.41	.21	***	15.8	***	.054	1.56		.017	-	8.47	***	.084	SS<SP>BU.NP
EV056	Lonely - desperate	-.56	1.3	.35	-.18	**	3.71	***	.099	1.53		.016	-	14.8	***	.139	SP.BU>SS
PA074	Punitive - antideviant	-.29	1.09	.45	.01		1.63	***	.037	1.26		.014		8.45	***	.084	SS<SP>NP.BU
RK032	Talkative - communicative	-.17	.63	.48	-.08		19.5	***	.065	1.3		.014	-	5.85	**	.06	BU>SS.NP
VA016	Verbally aggressive	-.73	1.15	.63	-.18	**	8.1	**	.028	1.1		.012	-	42.6	***	.317	SS<SP.BU>NP
ZS050	Ensuring future	-.23	1.01	.47	-.07		5.98		.021	2.81		.03		5.88	**	.06	SS<SP
GF106	Habit fixed	-.04	.96	.5	.06		1.57		.006	2.81		.03		5.3	**	.054	SS<SP
FG034	Fearful - security searching	0	.82	.56	-.32	***	3.84		.014	2.52		.027	-	5.38	**	.055	BU>SS.NP.SP
UG061	Weary - bored	-.12	.84	.51	-.08		3.69		.013	1.64		.018	-	5.01	**	.052	BU>SS.SP
WS039	Bold-sexually	-1.12	1.62	.62	-.04		3.81		.013	1.47		.016		155	***	.627	SS<BU<SP>NP
SZ047	Pugnacious - angry	-.16	1.02	.48	-.2	**	2.05		.007	1.49		.016	-	9.65	***	.095	SP>SS
ST023	Sexually driving force	-.19	1.06	.46	.03		1.72		.006	1.01		.011		4.5	**	.047	SS<SP
TG015	Loyally attached	-.39	1.41	.53	-.17	**	.45		.002	.81		.009	-	9.63	***	.095	SS.NP<SP
CG102	Charming - winning	-.05	.81	.57	-.11		1.25		.004	.75		.008	-	6.37	***	.065	SP>SS.NP
GA038	Emotionless - alexithymic	.02	.74	.55	-.02		.27		.001	.73		.008	-	3.97	**	.041	BU>SS.NP
SR078	Self-assured - robust	.04	.63	.61	.21	***	.88		.003	.66		.007		7.24	***	.073	BU>SS.SP
PU022	Polygamous - unfaithful	-.08	.7	.68	.05		3.41		.012	.62		.007	-	5.2	**	.053	SS<BU
IV008	Idealizing - deifying	-.19	1.07	.51	-.13		4.75		.017	.53		.006		6.33	***	.064	SS<SP>BU
TR024	Traditional - fixed on roles	-.02	.67	.55	-.1		.2		.001	.17		.002	-	6.1	***	.062	SS<SP.BU
DH073	Dependent - helpless	-.09	.64	.52	-.21	***	5.73		.02	.05		.001	-	5.52	**	.057	SP.BU>SS

		Descriptives T1-T2				Innersubject Over time			Between subject Subsample * T2-T1 differences				Between subject Subsample * T2-T1 distances			
Abbr.	Facet	m	SD	r	r age	F	p	ηp²	F	p	ηp²	Post hoc	F	p	ηp²	Post hoc
VE111	Understanding - sensitive	-.15	.58	.47	-.08	17.2	***	.058	4.76	**	.049	SP<BU	3.82		.04	
LE060	Passionate - emotional	-.23	.8	.59	-.17 **	24	***	.079	4.6	**	.048	SP<SS.BU	3.36		.035	
OF095	Obsessive - partner fixed	-.39	.91	.48	-.22 ***	52.3	***	.158	5.14	**	.053	SP<BU	2.91		.031	
SS028	Self-damaging	-.38	.84	.51	-.16 **	57.8	***	.172	4.14	**	.043	NP<BU	2.97		.031	
MK021	Preferring artistic - cultural	-.19	.84	.52	.1	13.8	***	.047	4.05	**	.042	NP<SS.SP.BU	2.46		.026	
EA052	Arrogant - overweening	-.09	.56	.57	-.19 **	8.04	**	.028	9.34	***	.092	SS.NP.BU<SP	3.21		.034	
NV001	In need for closeness and affiliation	-.28	1.09	.4	-.16 **	1.21	**	.035	9.6	***	.094	SS.NP<SP	2.26		.024	
DN082	Depressive - negativistic	-.11	.95	.57	-.23 ***	.74		.003	8.27	***	.082	NP<SS.SP.BU	2.51		.027	
MB004	Suspicious - insecurely attached	-.32	1.19	.64	-.28 ***	.21		.001	6.7	***	.068	SS.NP<SP.BU	1.62		.017	
UA005	Involuntarily adapted	-.01	.78	.48	-.12	.03		<.001	8.01	***	.08	SS.NP<SP.BU	1.47		.016	
GZ049	Happy - satisfied	-.61	1.42	.54	.16 **	1.98		.007	6.6	***	.067	SP<NP>BU<SS	1.4		.015	
HD063	Domineering - dominant	.02	.7	.53	-.13	.34		.001	4.16	**	.043	BU<NP.SP	3.49		.037	
PT019	Primitive - instinctively	-.19	.96	.53	-.1	4.78		.017	5.17	**	.053	SP<SS	3.17		.033	
BH086	Relationship questioning	-.18	.75	.66	-.25 ***	16.2	***	.055	2.3		.024	-	3.82		.04	
SH096	Beauty-loving - hyper aesthetic	-.15	.73	.6	-.14	11.3	***	.039	1.21		.013	-	3.82		.04	
UF108	Unconventional - liberal	-.23	.73	.67	.08	26.8	***	.088	2.26		.024		3.2		.034	
MG110	Materially generous	-.16	.73	.61	-.03	14.2	***	.048	.94		.01		3.11		.033	
SA007	Beauty conscious-attractive	-.34	.72	.61	-.14	6.07	***	.177	1.3		.014		2.83		.03	
RL094	Romantic-love illusionary	-.23	.9	.54	-.18 **	17.5	***	.059	.71		.008		2.1		.022	
BE109	Insulted - huffy	-.17	.77	.55	-.14	13.2	***	.045	.17		.002		2.02		.021	
EK020	Jealous - controlling	-.24	.8	.67	-.25 ***	11.3	***	.039	2.87		.03	-	.55		.006	
IA031	Intimacy-aversive	-.35	.7	.45	-.12	68.5	***	.197	.82		.009		.51		.006	
SE044	Sensible - emotional	-.14	.76	.61	-.2 ***	8.92	**	.031	.4		.004		2.02		.021	
LA092	Demanding lover's qualities	-.11	.72	.53	.02	6.8	**	.024	.04		<.001		1.43		.015	
AE115	Demanding attention	-.08	.51	.54	-.05	7.47	**	.026	.59		.006		1.04		.011	
RE064	Erotic - imaginative	-.15	.8	.43	-.02	9.69	**	.034	3.55		.037		.84		.009	
KU053	Disturbed communicating - misunderstood	-.21	.87	.51	-.15	7.95	**	.028	2.66		.028		.75		.008	
FW010	Amicable - appreciative	-.7	1.76	.36	0	9.9	**	.034	1.65		.018		.38		.004	
TA051	Tolerant - accepting	-.16	.86	.35	.07	5.94		.021	1.43		.015	-	3.15		.033	
KV040	Avoiding conflicts-closed	.02	.74	.39	.05	1.62		.006	1.01		.011		2.04		.022	
ER009	Emancipatory - radical	-.27	1.18	.45	-.07	.94		.003	.62		.007	-	3.05		.032	
VG026	Trustful - salvaged	.01	.72	.47	.13	.08		<.001	.64		.007		2.75		.029	
FH062	Caring - helpful	-.22	1.06	.49	-.05	.03		<.001	1.46		.016		3.33		.035	
VV101	Willing to reconcile - forgiving	-.14	.8	.5	.2 ***	4.48		.016	1.13		.012		1.15		.012	
VS114	Providing - securing	-.08	.9	.5	.03	0		<.001	1.07		.011		2.61		.028	

Abbr.	Facet	Descriptives T1-T2				Innersubject Over time			Between subject Subsample * T2-T1 differences				Between subject Subsample * T2-T1 distances			
		m	SD	r	r age	F	p	ηp^2	F	p	ηp^2	Post hoc	F	p	ηp^2	Post hoc
SU002	Sexually unsatisfied	-.12	.89	.5	-.01	5.14		.018	2.26		.024		2.5		.026	
DB012	Preferring distance	-.16	1.05	.51	.21 ***	.03		<.001	.2		.002	-	3.72		.039	
PD099	Perverse - deviant	-.04	.88	.51	.05	.58		.002	2.99		.032		3.08		.032	
KW080	Church-traditional valuing	-.02	.91	.52	-.24 ***	.2		.001	.46		.005		3.64		.038	
PV098	Varying in partner type	-.1	1.17	.52	.04	1.81		.006	1.45		.016		2.74		.029	
BV105	Protective - defending	-.05	.8	.52	-.07	1.17		.004	.43		.005		2.67		.028	
WI070	Daring - imposing	-.06	.7	.52	-.13	2.27		.008	.29		.003		1.02		.011	
FD055	Leadership quality - dominant	-.01	.63	.53	.03	.02		<.001	.15		.002		2.23		.024	
US035	Independent - self-determined	-.03	.73	.54	0	.35		.001	.82		.009		.53		.006	
MP027	Macho - boastful	-.06	.79	.55	-.25 ***	1.38		.005	1.78		.019		2.29		.024	
SF089	Eager to look good	-.06	.73	.55	-.25 ***	2.19		.008	.68		.007		1.9		.02	
LB072	Qualified as lover	.04	.64	.57	-.03	1		.004	2.86		.03		1.31		.014	
LK041	Motivated to achieve - carrier-orientated	-.23	10	.58	-.25 ***	6.4		.022	1.64		.017		2.67		.028	
LF097	Fixed on lover's qualities	-.09	.75	.58	-.1	3.86		.014	.06		.001		1.27		.014	
PB025	Phlegmatic - comfortable	-.02	.7	.58	-.12	.24		.001	.16		.002		.64		.007	
RN068	Vindictive - vile	-.06	.92	.59	-.22 ***	1.2		.004	.49		.005		2.11		.022	
WA066	Workaholic	.08	.68	.59	.03	3.43		.012	.55		.006		3.77		.039	
KA011	Physically abusive	-.09	.7	.59	-.06	4.76		.017	1.65		.018		2.71		.029	
AS079	Adventurous - sporty	-.08	.74	.59	-.1	3.4		.012	3.09		.033		1.02		.011	
SG065	Shy - inhibited	-.03	.76	.6	-.17 **	.48		.002	1.68		.018		2.51		.027	
GE090	Demanding *conscientiousness*	.03	.53	.6	-.13	1.18		.004	.95		.01		2		.021	
ES018	Esoteric - spiritual	-.16	.91	.64	.04	1.09		.004	2.82		.03		2.44		.026	
GH103	Good-natured – humorous - agreeable	-.06	.58	.65	.06	2.8		.01	3.08		.032		3.78		.039	
SV084	Fearing sexual failure	-.18	.89	.65	.02	3.47		.012	1.37		.015		.29		.003	
SN058	Stressed - nervous	.07	.66	.67	-.22 ***	3.35		.012	1.89		.02		3.54		.037	
DS014	Being a rough sexist	-.08	.73	.67	-.32 ***	3.3		.012	.77		.008		1.78		.019	
IE112	Intercultural - exotic loving	-.1	.99	.67	.17 **	.44		.002	.14		.001		1.09		.012	
PS083	Promiscuous - partner changing	.04	.7	.7	0	1.15		.004	.5		.005		1.32		.014	

Notes. *Facets of the Relationship- and Attachment-related Personality Inventory: Facets are ordered by significant effects, e.g., facets which decrease or increase differently between the subsamples and which additionally change more for some subsamples than for others are in the first category on top of the table. The category on the bottom of the table is ordered by retest-correlation over the whole sample (column r).*

Results of MANOVAS on differences and distances between T1 and T2 as dependent variables. Difference: T2 minus T1. Distance: number of the difference.

df1=3, df2=276. m and SD of mean-level-changes between T1 and T2. r age: correlation of T1 scale value with age (Pearson). r: correlation between T1 and T2 (Pearson).

*Post hoc: results of significant differences between the 4 subsamples. SS: still-singles, n=74. NP: new-partners, n=40. SP: still-partners, n=102. BU: break-ups, n=64. T1: Time 1. T2: Time 2, four years later. **p<.01. ***p<.001*

3.3.4. Figures for significant relationship-related mean-level changes

Figure 11, *Figure 12* and *Figure 13* visualize the T2-T1 facet changes which were described in the previous section and which significantly differed between subsamples[17].

Figure 11 shows seven facets which have increased or decreased differently in different subsamples and additionally decreased over time overall, e.g., *self-damaging* (SS028) significantly decreases over the course of the four years and less does this for *break-ups* than for *new-partners*: therefore, ochre-colored lines for *break-ups* and *new-partners* do not overlap on the x-axis and the vertical ochre-colored lines represent the average shift between T1 and T2 measurement over the whole sample which is significantly different from zero.

Figure 11. Part I. Facets which differed in shifts over time between the subsamples (n=280)

Notes. *Facets of relationship-related personality which significantly differed in their T2-T1 differences for different relationship surroundings. Error bars with 95% confidence interval around the subsample's averages. Vertical lines: average of changes for the whole sample, only presented if significantly different from zero. Values on the x-axis: reflect the size of the changes scored in scale points; If T2=T1: x=0. T1:Time 1. T2:Time 2. Example: NV001 decreased over time overall and more for new-partners than for still-partners.*

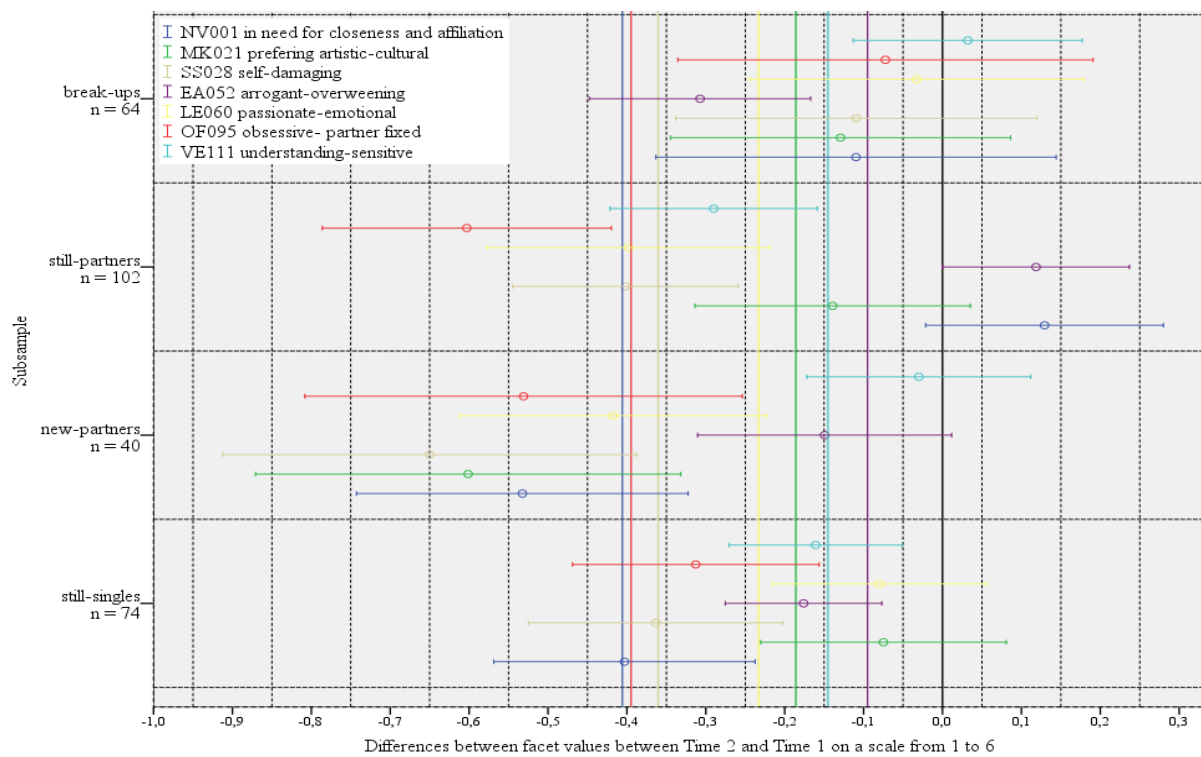[17] Appendix II.5. contains figures for changes in T2-T1 facet distances.

Figure 12. Part II. Facets which differed in shifts over time between the subsamples (n=280)

Notes. Facets of relationship-related personality which significantly differed in their T2-T1 differences for different relationship surroundings. Error bars with 95% confidence interval around the subsample's averages. Vertical lines: average of changes for the whole sample, only presented if significantly different from zero. Values on the x-axis: reflect the size of the changes scored in scale points; If T2=T1: x=0. T1:Time 1. T2:Time Example: BS113 decreased more for new-partners than for still-partners.
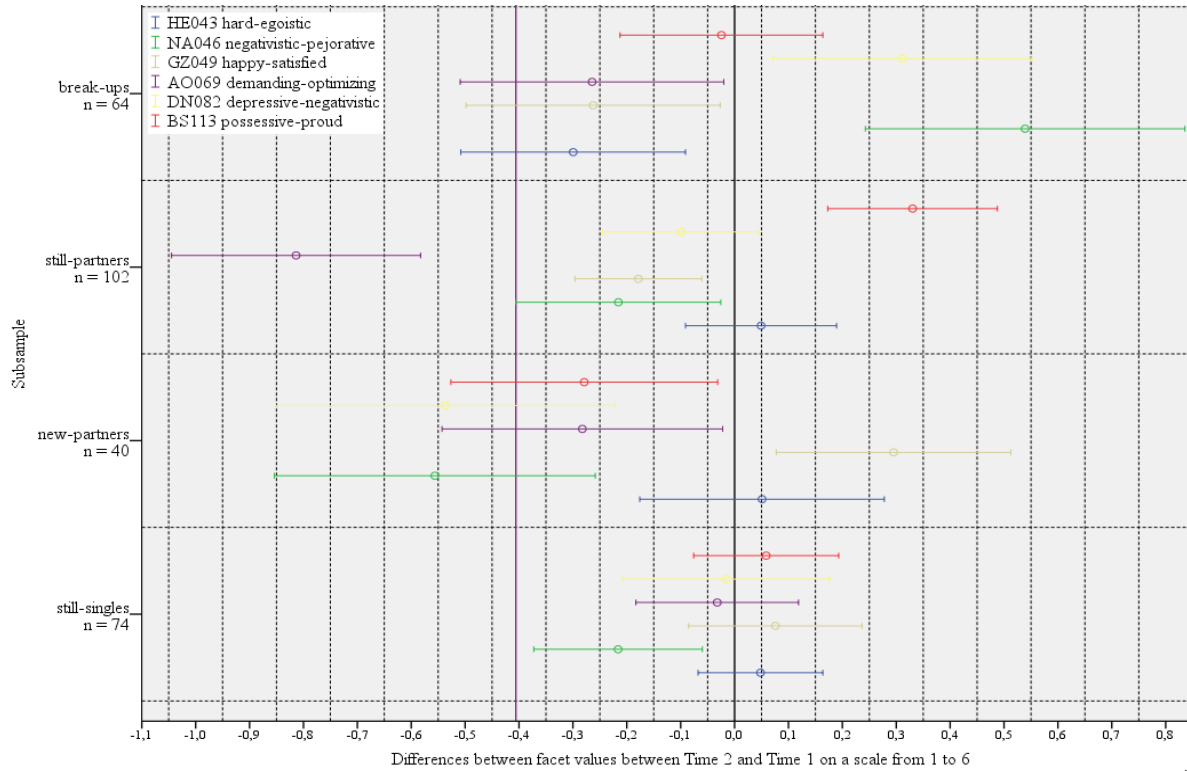

Figure 13. Part III. Facets which differed in shifts over time between the subsamples (n=280)

Notes. Facets of relationship-related personality which significantly differed in their T2-T1 differences for different relationship surroundings. Error bars with 95% confidence interval around the subsample's averages. Values on the x-axis: reflect the size of the changes scored in scale points; If T2=T1: x=0. T1:Time 1. T2:Time 2. Example: MB004 and UA005 decreased more for new-partners than still-partners or break-ups.
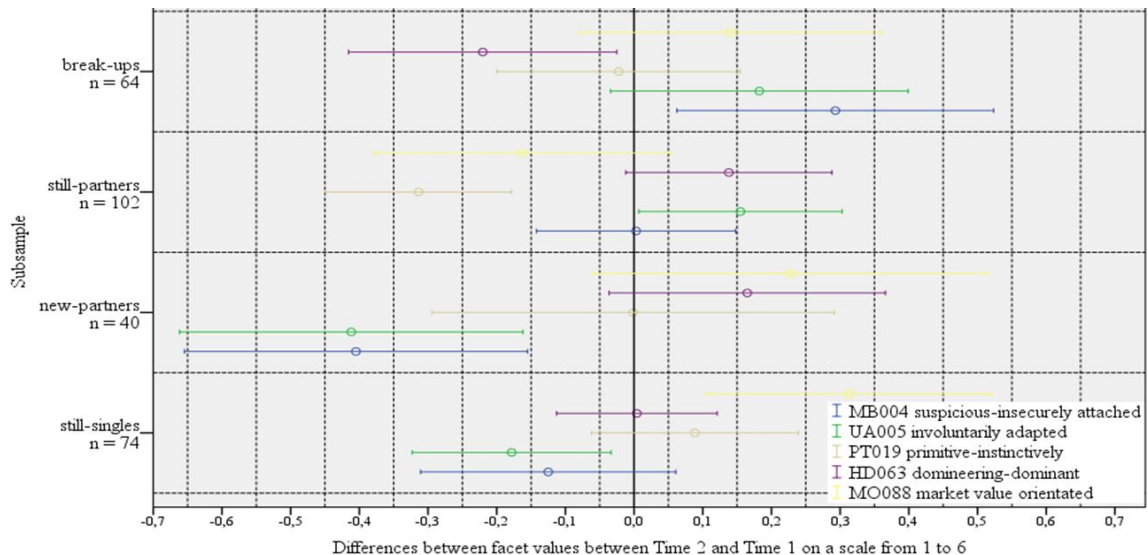
3.4. Discussion

3.4.1. Main results

Individual retest-stability

The profile correlation between the early measurements and the measurements four years later was high on average and for most subjects. Current partnerships (at T1 and/or T2) negatively affected the personal measurement stability, while the number of relationships before T1 or between T1 and T2 had no influence. Thereby, *hypothesis 1b,* that retest-stability correlated with relationship experience, was partly supported. Furthermore, *hypothesis 1a,* that retest-stability is slightly associated with age, but not with gender, was confirmed. As assumed by *hypothesis 1c*, the retest-stability also significantly differed between subsamples as follows: *still-singles > new-partners = still-partners > break-ups*.

Changes in facet values over time

Figure 14 pictures how many of the 90 facets have been affected by different kind of changes over time and how these changes overlap. To summarize, 35/90 facets (38.9%) were robust against mean-level-changes and subsample differences in alterations. The other 55 facets were affected by various T1-T2 changes:

- All 90 facets significantly changed over time and, as hypothesized (*3*), for 28/90 facets (31.1%), these T2-T1-*distances* significantly differed between the subsamples;
- 41/90 facets (45.6%) were affected by significant mean-level-changes:
  - As hypothesized *(2a),* 30/90 facets (33.3%) significantly decreased and 1/90 facet *(*1.1%), increased over time;
  - As hypothesized *(2b),* 18/90 (20.0%) facets significantly decreased or increased more or less depending on the subsample.

As a further indication for *hypothesis 2a,* that facets increase or decrease over time*,* 36/90 facets (40.0%) significantly correlated with age, 7/90 (7.8%) positively and 29/90 (32.2%) negatively.
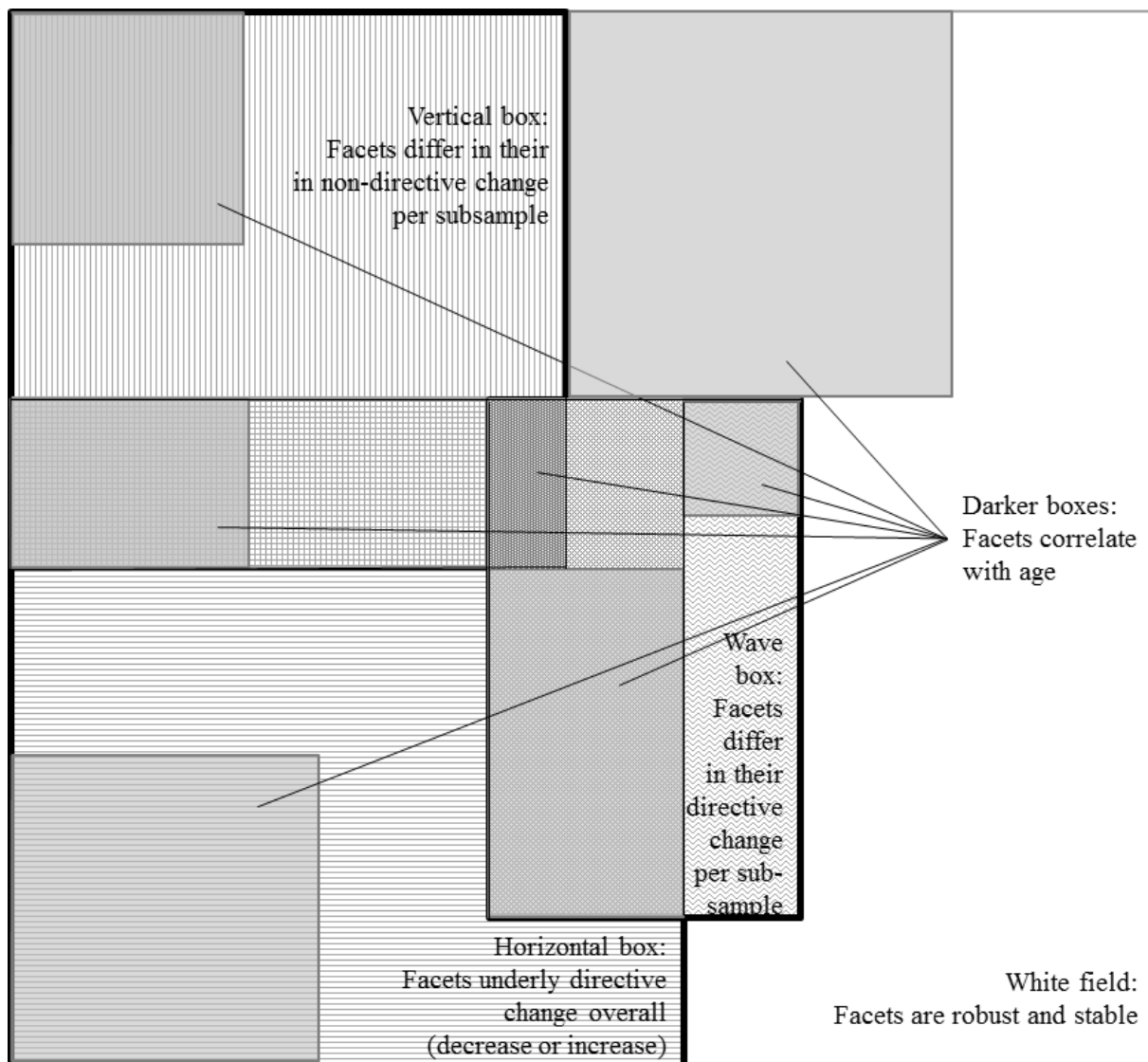
Figure 14. Proportions (%) of the 90 facets which were affected by various changes over the

four years (n=280).

Notes. The larger an area, the more the corresponding kind of change affected the facets.
Where patterns overlap, all corresponding effects affected facets.
Wave box: Confirming hypothesis 2b. Horizontal box: Confirming hypothesis 2a.
Vertical box: Confirming hypothesis 3a. Darker boxes: Further indication for hypothesis 2a.
White field: Corresponding hypotheses all were all denied.
Directive changes: Time 2-Time 1 facet differences. Non-directive changes: Time 2-Time 1 facet distances.

3.4.2. Conclusion

3.2.1.1. Effects of time, age and relationship status

Facets of *relationship-related personality* were all at least moderately stable over four

years ($0.3 < r < 0.7$). These retest-stabilities were lower than the ones found for domain scales

in the previous one-year-study on the BB-PI (Wunderlich, 2011), and similar to those for

similar time spans for other self-assessed relationship-related dimensions, such as facets of

agreeableness from the common German test version for the *Big Five* (Ostendorf & Angleitner, 2003). This finding indicates acceptable general mid-term retest-reliability for the facets of the BB-PI.

As in corresponding reviews (Roberts & DelVecchio, 2000), the present results suggest that *relationship-related personality* becomes more stable with age. Nonetheless, the individual differences between T1 and T2 were twice as large as the inter-individual differences associated with age. Many of the facets which were related to inverse *emotional stability*, negatively (SR078, SN058, SE044, EK020, BH086, NV001, SS028, OF095, MB004), while many conflict- and *agreeableness*-related facets positively correlated with age (VV101, RN068, HG037, VA016, SZ047, HE043, NA046, AO069). This finding is congruent with preceding research results that *agreeableness* increases and *emotional stability* decreases with age (Roberts, Walton, & Viechtbauer, 2006). The BB-PI scales measure a large number of negatively connoted and extreme aspects concerning the mating personality. Since people tend to functionally adapt to an existing environment and the environment tends to 'grind sharp edges', both over time, the facet values may have decreased rather than increased (34.4% versus 1.1% of the 90 facets).

The differences in retest-stabilities between subsamples indicated that changes in *relationship-related personality traits* are associated and interact with current relationships. In particular, having a relationship when completing the questionnaire (at both times) seems to have a large impact on how the corresponding questions are answered, while relationships in between the surveys did not destabilize the measurements. Possibly, individuals believe that the tendencies of their own thinking and behavior which they experience in a current partnership are particularly representative for their own personality: the value and generalizability of this self-perception is probably overestimated in form of an error of reasoning, based on the availability, presence, emotional relevance and exclusivity (in case of monogamy) of the information. A possible reason why *new-partners* were more stable in their traits than the *break-ups* may be that the influence a relationship has on personality self-assessment rises with its duration: The relationships that the *break-ups* left were on average longer than those entered by *new-partners* (max. four years in the current study).

3.2.1.2. *Subsample differences in facet changes*

For 40/90 facet scales (44.4%), T2-T1 changes differed between subsample in a way which indicates that the corresponding facets are probably confounded with current relationships, .i.e., that the traits might change when entering a new or living in a relationship and

altering back when leaving it. This applied for very heterogeneous contents but one pattern can be observed: particularly facets linked to *tendencies to emotionally react negatively or positively*, were related to varying changes over the different relationship surroundings (see *Table 15*). For instance, MB004 (suspicious - insecurely attached) decreased more for *new-partners* and *still-singles* than for *still-partners* or *break-ups (see Figure 13).* These results expand and replicate former findings concerning the *break-up*-related changeability of attachment styles (Kirkpatrick & Hazan, 1994), to the whole range of *relationship-related personality* and to different status changes: a general relationship-related adaptiveness of relationship-related personality traits was supported. However, the assumption from the previous study on the BB-PI (Wunderlich, 2011) that facets, linked to submissive-protective domains, were less stable than those, linked to dominant offensive ones, could not be confirmed.

3.4.3. Possible implications for predicting relationship quality

Some of the adaptive facets are probably confounded with the RQ: some may be influenced by it, while some may influence it. RQ-influenced facets would rather be valid for a specific relationship than for different relationships, and appear as an indirect measurement of RQ itself. Therefore, these might not be appropriate to predict the future RQ with a new partner. For instance, somebody could assess oneself as insecure attached because their current relationship is troublesome and, thereby, makes them feel insecure, while a potential new relationship, which might be more securing, could, thereby, be accompanied with a higher quality. The adaptiveness of these especially predictive scales represents a considerable problem for an application in mating contexts and could be one of the main reasons why online dating based on partner matching has not yet yielded higher RQ (Cacioppo, Cacioppo, Gonzaga, Ogburn, & VanderWeele, 2013): Since RQ models are typically developed on partners who already influenced and adapted to one another, they might not be transferable to singles who not yet know one another. Although *relationship-related personality traits* are less consistent over different relationship environments than *general personality*, they were found to be additionally important when predicting RQ (see *Study I: Long-term Prediction of Relationship Quality with Machine Learning by Personality Traits*). Therefore, approaches to handle the facet changes are desirable: Two approaches are described in the present section.

3.4.3.1. Handling mean-level changes

Since facets decrease or increase over time, their calculated risk or bonus effects on RQ may become stronger or weaker. A method to handle this in the context of predicting RQ,

could be to determine patterns of changes in norming samples and, based on that, to correct the facet values of tested subjects by the expected mean differences. The present findings support that varying mean-level-differences for different relationship surroundings occur. Considering these may contribute to transferring models from one subsample more validly to the other: from couples to singles preferably.

Implications for predicting RQ are explained using three examples:

(1) The scores for the facet *demanding - optimizing* ranked among the most important negative partner effects predicting RQ (*Study I*: *Table 9*, AO060). In the present study, AO060 decreased over time with an effect size larger than one SD (AO060: $\eta p^2=.139$, p<.001, n=280). This indicates that the potential negative impact on RQ might reduce over time as well, especially since the reduction was even stronger for *still-partners* than for other subsamples.

(2) The scores for the facet *happy - satisfied* increased more for *new-partners* than for *still-partners* or *break-ups*, i.e., *new-partner's* RQ might profit from a boost which later vanishes again. Therefore, mid-term actor effects may be underestimated for singles in mating contexts when the predictions are based on couple's data (S*tudy I: Table 9*, GZ049).

(3) The scores for the facet *understanding - sensitive* decreases over time, more for *still-partners* than for *break-ups*. Therefore, the positive actor effect on RQ (*Study I*: *Table 9*, VE111) may subside in a partnership over time and would be stronger for subjects in a mating context. Another explanation for the differing changes might be that partners decrease more in VE111 because they already understand one another, while *break-ups* sustain this trait to better understand potential new partners.

3.4.3.2. *Handling retest-stability*

Pre-selecting facets, which are sufficiently immutable over time and relationship surroundings, could ensure the transferability of models from a couple's context to mating contexts. Higher retest-correlations of facets indicate that aspects of the environment less and aspects of the person more influence the corresponding trait. Especially retest-reliable facets might be better suited as long-term-predictors, since their values remain available over time. The calculative assumptions to explain this consideration is presented in Table 16: using a sample of approximately 200 subjects (such as in *Study I)*, an effect is required to at least potentially predict 1.44% of the variance of an outcome in order to become detectable, i.e. to become significant (r>.12). This is only possible for facets with retest-stabilities of r>.30 -

provided the stable part of the variance is also the actually predictive variance. In the present study, retest-stabilities for 23/90 facets (25.6%) fell below this criterion in at least one subsample (see *Table 12*).

Table 16: Hypothetical scenario –explained variance of an outcome over time by predictors with varying retest-reliabilities and predictive validities

r1 assumed four-year-retest-correlation of a characteristic	r2 cross-sectional correlation of a characteristic with an outcome (high. middle. low)	r^2: $r1^2 \cdot r2^2 \cdot 100\%$. Four year explained variance of the outcome after 4 years in %
	0.7	39.69
0.9	0.5	20.25
	0.3	7.29
	0.7	24.01
0.7	0.5	12.25
	0.3	4.41
	0.7	24.01
0.5	0.5	6.25
	0.3	2.25
	0.7	7.84
0.4	0.5	4.00
	0.3	1.44
	0.7	4.41
0.3	0.5	2.25
	0.3	0.81

Facets which differ in their retest-stabilities between subsamples, may to a certain extent adapt to different relationship surroundings over time. Implications for predicting RQ are explained using three examples:

(1) Most facets linked to sexual preferences (8 of 11) were especially robust against current relationship surroundings. These findings indicate that these facets may be applicable for matching models from this perspective. They could predict cross-validated 18% of the variance from *sexual satisfaction* over time (*Study I*: *Table 7*). To conclude, this model may probably also be valid for matching singles in a mating context.

(2) Scores for the facet *negativistic-pejorative* changed so largely for *break-ups* (NA046, retest-stability: r(278)=.22), that the large actor effect (*Study I,* RQ- NA046-correlation: r(190)=.57) would only explain 1.57% variance of the RQ for a new partnership (r^2=0.22*0.22 *0.57*0.57*100% = 1.57%). Thus, for this purpose, NA046 might be nearly useless.

(3) Scores of the facet *sexually unsatisfied* were less retest-reliable than scores of the facet *suspicious-insecurely attached (*SU002: r(278)=.50, MB004: r(278)=.64), while CC_ll was among the most important predictors for both *(Study I: Table 9)*. Since the facet values for both partners are more likely to change for SU002 than for MB004,

the RQ-associations with MB004 may also be more stable over time. Therefore, despite the similar variable importance within the partner model, MB004 may be more appropriate as long-term-predictor than SU002.

To raise an objection to the presented idea, some traits could be unstable over time, but nonetheless, predict future success in another way, rather than by their consistent measurability in the future. Using the example of the well-known Marshmallow Test (Mischel, et al., 2010), the potential operating principle is explained: children who proofed enough willpower and trust to wait with eating the first available treat (marshmallow), in order to get a second one later, turned out to make more successful careers in adult age. Of course, the test result itself cannot be replicated using the same test in a later age, though early results still predict the later life outcome. Similarly, BB-PI facets may be partly instable over time or during relationships, but still evolve to be valid predictors for the RQ of future partnerships, i.e., partners might develop certain skills or insights during a partnership which can be activated again by future relationships. Even or even especially in an unhappy commitment the parties might improve their understanding and handling of a partners personal or gender-related specificities. Correspondingly, longitudinal predictions of RQ based on the BB-PI differed from cross-sectional prognoses (*Study VI:Table 24*). This finding suggests that the BB-PI measures facets which work as described. Future studies have to show if this is true and if preselecting robust facets is helpful for predicting.

3.4.3. Limitations and outlook

The results of the present study gave a first overview and impression of *personality traits* which probably are confounded with relationship status and RQ.

The subsample of *new-partners* was exceedingly small in the present study, though this is the target group in a mating contexts in which the BB-PI ideally should be applicable. On a representative number of subjects, patterns of changes and stabilities should be determined to more accurately evaluate which and to what extent the different facets adapt to a new relationship. Additionally, the question concerning partnership-overarching traits should be addressed in more detail, using a larger sample of subjects who change their partner. Future studies should employ time series analyses to determine which relationship-related experiences exactly (conflicts, attraction, e.t.c.) correspond with changes in personality. This could be a step to find out, how re-learning certain risk behavior or developing social skills which protect RQ can be accelerated.

The applicableness of the present results for mating contexts is also restricted for another reason: influences of trait changes on predicting RQ could only be suspected but not be proven. Future work must directly examine, to what extent *personality traits* which are measured during one partnership are able to predict the RQ of a subsequent relationship. If the prediction quality reduces in this case, the presented approach to correct for expected mean-level changes can be tested and evaluated as a method to handle this.

4. *Study III: Psychometric Evaluation of Scales for Preferred and Own Personality Traits*

Abstract

The present study aims to develop and validate a questionnaire for self-assessment of general and relationship-related personality, which would be highly applicable in real-life mating contexts and have a maximum potential predictive validity for relationship quality. Scales were designed using a combination of rational and correlation-based test construction. In order to measure similarity as well as ideal-real fitting of partners, the test contains and parallelizes the same scales for the own and preferred partner characteristics. In heterogeneous samples of 423 and 250 subjects, acceptable retest-reliabilities, internal consistencies, typical gender and nationality differences, as well as convergent, discriminant, internal, and predictive validities were observed.

Predictive validities for actor, partner, fitting, and similarity scales with RQ were presented for 192 participating partners (96 couples). A novel scaling to score different constellations of the partners' values were included (Combination Counts): *Study I*'s findings concerning similarities that shared weaknesses often negatively correlated and shared strengths often positively correlate correlated with RQ were replicated. The explorative analysis of predictive patterns for fittings indicates an activating function of preferences: if the actor highly preferred a trait, high partner values in this trait were often beneficial and low ones restraining, while the partners values rarely had an effect on relationship quality if the actor did not highly prefer this trait.

Keywords: *personality traits*, *relationship satisfaction*, partner preferences, close relationships, partner similarity, psychometric validation, test validation

4.1.   Additional background[18]

Matching tests should be able to forecast how happy two individuals will become in a future partnership with one another. Relating thereto, some special requirements, which go beyond the common test criteria (Rost, 2004), need to be fulfilled and, in this section, are derived from previous research.

4.1.1. Relationship quality, preferred and own personality

Preceding research suggested that compatibilities of ideal and real romantic partner are highly associated with RQ: (1) longitudinally (Fletcher, Simpson, & Thomas, 2000), (2) cross-culturally (Gerdvilyte & Abhyankar, 2010; Ruvolo & Veroff, 1997)**,** and (3) even slightly more than partner similarities (Zentner, 2005). Hence, to be maximally effective, a matching test should enable dyadic ideal-real *fittings* derived from self-assessments in addition to measuring partner similarities: e.g., in practice, the popular dating website OkCupid - which was invented by mathematicians - matches potential partners using a *fitting* score (Humor Rainbow, 2015). Unfortunately, the algorithm has not yet been validated with regard to RQ.

Studies have shown that people generally prefer partners similar to them (Botwin, Buss, & Shackelford, 2006; Regan, Levin, Sprecher, Christopher, & Gate, 2000). Due to this moderately correlating constructs, many personality tests mix contents of preferred and own traits within the same scales, even if arising scales might not enable an optimal matching. A questionnaire employing a promising alternative approach is the Gießen-Test (Brähler & Brähler, 1993), which refers to all types of relationships. Preferences for characteristics in the counterpart and scores for own characteristics are completely separated from one another while corresponding scales are parallelized. Until now, no questionnaire that specializes on romantic partnerships exclusively has been published. The present study has developed such a test.

4.1.2. Combinations Counts

In *Study I*, partner similarities measured with CCs were examined for their effect on RQ: shared weaknesses affected RQ negatively and shared strengths positively (see *2.3.1.2. Predictive scale validity*). The present study followed up *study I* by testing predictive validi-

---

[18] *Preceding background sections which are additionally relevant for this study:*
2.1.1. Actor-, partner- and similarity effects
2.1.2. Trait-specific effect of similarities
2.1.3. Effects of relationship-related and general personality

ties of CCs for similarities with another personality test, going on to evaluate the CCs of fittings for the first time. Fittings might affect RQ differently when compared to similarities: e.g., it is possible that self-assessed own partner characteristics only affect RQ when the actor prefers this characteristic in a partner.

4.1.3. Developing process

Andresen developed an early short version of the current YOUME-L, entitled 'Wish Partner- Real Partner- Lists'. The construction sample was congruent with the cross-sectional *Stern* dataset (see 1.2.2 *Data collection*): the survey measured self-assessed preferred partner characteristics, but own characteristics were assessed by the current partner and not by oneself. The items moderately correlated with corresponding BB-PI scales, showing external validity (Andresen, 2012). On the same dataset, Smolic (2012) developed a factor analytical scaling with eight factors and offered a first impression of predictive validity: the partner-assessed own characteristics and their fittings to the self-assessed partner preferences mainly moderately correlated with RQ.

The early YOUME-L had some two shortcomings which were addressed and optimized in the current study:

- In addition to *general personality*, *relationship-related personality* characteristics were shown to be important when predicting RQ (see *Study I: Long-term Prediction of Relationship Quality with Machine Learning by Personality Traits*). The early version of the YOUME-L does not measure some of these attitudes and attachment characteristics, such as trust and communicativeness, which were found to be relevantly associated with RQ in existing research studies, such as the one by Shaver (1992).

- The previous scaling of Smolic (2012) integrated domain-like scales linked to *openness*, *conscientiousness*, and *extraversion*: these alone would not allow a more differentiated matching. To efficiently find and combine effects on RQ, many different homogenous scales for various characteristics are required - not only a few domains which integrate many correlated facets. Based on such a test, exclusively scales which are relevantly involved in predicting RQ for couple's could be economically selected and be later applied in a more goal-oriented way in real-life mating contexts.

- On the six-point *Wish-Scales*, the lowest point states 'very unimportant' and the highest states 'very important'. Given the comparably low scale variances (Andresen, 2012), this scaling may grade more to better differentiate, e.g., by including 'unwanted' to the low and 'indispensable' to the high end of the scale.

4.1.4. The present study

To enable dyadic ideal-real fitting in addition to similarity scales for potential partners, a test which separates and synchronizes homogenous scales for preferred partner (*Wish-Scales*) and own characteristics (*Self-Scales*) was designed. Various aspects of relationship-related and *general personality,* as well as of attraction, were implemented by rational test construction and combined with correlation-based methods. Thereby, the short version of the YOUME-L was optimized in multiple ways. Common test validation was conducted for a long and short test version on a heterogeneous sample of 423 (additional 250 subjects for retest-reliability).

The main focus was predictive validity which were evaluated for the RQ of 192 partners:

(1) Correlations were compared with those from previous research findings on RQ-related characteristics;

(2) Patterns of actor, partner, similarity and fitting effects were determined overall, and for different subsamples of gender and culture;

(3) Patterns of effects for the novel CCs with RQ were replicated for similarities and explored for fittings.

4.2. Methods

4.2.1. Operationalization

4.2.1.1. Personality

The YOUME-L contains two lists of the same item contents: (1) participants' assessed to which extent characteristics apply to themselves and (2) to which extent the characteristics should apply to a potential romantic partner. Both lists use ordinal scales:

- *Wish-Scales*: 0: 'unwanted', 1: 'very unimportant', 2: 'unimportant', 3: 'rather unimportant', 4: 'rather important', 5: 'important', 6: 'very important', 7: 'indispensable'.

- *Self-Scales*: 1: 'very false', 2: 'false', 3: 'partly-partly', 4: 'rather true', 5: 'true', 6: 'very true'.

For example, participants rated to which degree *gregariousness* normally applies to their own personality and how important gregariousness is to them in an ideal mate. This way, the lists directly, economically and comparably measure preferred and own characteristics, as suggested in the corresponding research (Botwin, Buss, & Shackelford, 2006). The earlier short version of the YOUME-L (Andresen, WPRP-L, 2012) contains 2x40 items. 2x34

were added: The long advanced version contains 2x74 items and, additionally, used 'indispensable' and 'unwanted' as answer categories for *Wish-Scales*.

The individually perceived importance of different life areas and the satisfaction within these was measured by the FLZ M (Henrich & Herschbach, 2000).

4.2.1.2. Fittings and similarities

Similarity and fitting scales were calculated by the following formulas using the items that belong to one facet (i: number of items).

$$SimilarityWishFacet = \sum_{k=1}^{Facet\ i} (|Actor\ Item\ Wish\ k - Partner\ Item\ Wish\ k|)$$

$$SimilaritySelfFacet = \sum_{k=1}^{Facet\ i} (|Actor\ Item\ Self\ k - Partner\ Item\ Self\ k|)$$

$$FittingP1P2Facet = \sum_{k=1}^{Facet\ i} (Actor\ Item\ Wish\ k\ /\ Partner\ Item\ Self\ k)$$

$$FittingP2P1Facet \sum_{k=1}^{Facet\ i} (Partner\ Item\ Wish\ k\ /\ Actor\ Item\ Self\ k)$$

While distances were applied for similarities, divisions were applied for fittings. As an example, the scales for *physical attractiveness* contains five item contents (PAT, i=5). To calculate 'SimWish PAT', the five distances between the item values of both partners were summed up. To calculate FitP1P2 PAT', the five divisions for the item values of the actors' preferences divided by the partners own characteristics were totaled.

CCs for similarities were calculated, as described in *Study I* (see *Time one similarities*). In the present study, the CCs were also calculated for all possible combinations of high and low values for fittings. Both patterns to categorize high and low values correspond with the right-sided value distributions: Regarding *Wish-Scales*, low values refer to 0, 1, 2 or 3, while high values refer to 5, 6 or 7; Regarding *Self-Scales*, low values refer 1, 2 or 3, while high values refer to 5 or 6.

Additionally, sum scores for all scaling types, as well as profile correlations, were calculated for similarities and fittings.

4.2.1.3. Relationship quality

RQ was tested by the same items described in *Study I* (see *2.2.1.3 Time two relationship quality*) - with the difference that the short (not the long) version of the FLZ was applied, the so-called FLZ M (Henrich & Herschbach, 2000). This version measures general *satisfaction* in sexuality and partnership with one item only. Data on RQ was completely available

for only the three items concerning *conflicts* (intensity, frequency, and duration), as well as for the item about *perceived constrictions due to the current partnership*. 159 of the 192 subjects additionally answered the item from the FLZ M. However, due to temporary technical difficulties, only 84 of the 192 subjects provided details about their *harmony* in different partnership domains and about their *separation indents*. After z-standardizing and polarizing item values into the same direction, all available values per person were used to calculate the RQ average, which represents *RQ overall*:

RQ overall = MEAN(z *invers conflicts, invers z restriction, z Satisfaction with partnership/sexuality, z item for general relationship satisfaction, z reversed separation indents, z partnership harmony).*

4.2.2. Samples

Two datasets were examined: (1) a CS with the long version of the YOUME-L for test construction and evaluation of internal and predictive validity, and (2) a LS with the short version of the YOUME-L for evaluating the retest-reliability of *Wish-Scales*.

(1) The cross-sectional sample consisted of 440 adults with above-average levels of education. 17 participants were treated as dropouts due to too many missing values, based on an outlier analysis or answer variance SV<0.15, leaving a sample of N=423. The average age was m=31.8 with SD=11.6 (Range: 16-75). 237 participants were female (56.0%), 148 were male (35.0%), 38 stated no gender (9.00%), 52 were Persian (12.3%), 371 were German (87.7%), 123 were single (29.1%), 300 were committed (70.9%). 192 of these attended the study with their partners (44.9%) and were used for analyzing predictive validities: A sample description is placed in *Study IV* (see *5.2.2 Sample).* 149 participants had a university degree (35.2%), 169 had a high-school diploma (*Abitur* in Germany) (40.0%), 59 had finished secondary education (13.9%), 10 had a lower level of education (2.4%) and 36 did not state their level of education (8.5%). No variable was missing more than 38 values.

(2) From the 250 attendees who participated at T1 and T2, the average age was m=41 with SD=12 (Range: 18-72). 133 were female (53.2%), 117 were male (46.8%), 100 were single (40%), and 150 were committed (60%) at T1. At T2, 80 participants had not changed their relationship status (32.0%). 137 participants had a university degree (54.8%), 42 had a high-school diploma (16.8%), 63 had finished secondary education (25.2%), and eight participants had no educational qualification (3.2%).

4.2.3. Procedure

4.2.3.1. Scale construction

The goal was to construct homogenous facet scales with three to eight items for both own and preferred characteristics. Additionally, larger, more heterogeneous domain scales associated with the *Big Five* were aggregated.

Items were clustered into scales: primarily based on rational considerations, aided by a series of exploratory factor analyses, computed with diverse extraction and rotation criteria based on the sample of 423 subjects. Principal Component Analyses were run with:

- 5, 8, 10, 13 fix factors and the factor number for the scenario of eigenvalue > 1 (Kaiser-Guttman-criterion),
- varimax rotation and oblinque rotation,
- Wish-Items only (1), Self-Items only (2), both (3).

A factor was extracted and validated when it reoccurred in the main components, and/or when its content was presented and replicated in previous research about personality or close relationships. Items mostly were allocated to the factor they were correlated the highest with. All scales should show: (1) acceptable factorial validity, (2) high part-whole corrected item-scale correlations (r >0.3), and (3) high internal consistency (CA>0.6). Based on these criteria, six items were excluded from the scaling (items: *mechanical skills, cuteness, helplessness, jealousy, non-smoker, romantic tendencies*).

4.2.3.2. Test validation

Common test validation was performed on the 423 individuals and included descriptive statistics, internal consistency, correlations of scales with one another (Pearson) an MANOVAS for detecting subsample differences in scales. Last mentioned were run for the following subsamples as independent variable: German singles vs. German partners; female partners vs. male partners of the same couple; Persian partners vs. German partners. The subsample of German partners was parallelized to the Persian one concerning relationship duration, age, education, number of children and living status[19].

For predictive validity, all actor, partner, fitting and similarity scales were correlated with the *RQ overall* (Pearson) - once for the whole sample of 192 partners and once for every subsample of nationality and gender (i.e., for 140 Germans, 52 Persians, 96 men and 96 women).

---

[19] Appendix III.5. presents the descriptive statistics of the YOUME-L scales for all relevant subsamples.

4.3. Results

4.3.1. Scaling

Table 17 presents contents of items and scales for both version of the YOUME-L, including the corresponding literature constructs for allocation. The names of the scales were chosen to match the well-known constructs on which contents were partially based upon in the rational classification. Roughly corresponding to the *Big Five* (Ostendorf & Angleitner, 2003)*,* five domain scales were integrated. These scales mostly subsume more items than the facet scales, e.g., *conscientiousness* mainly contains the items of the scales *responsibility*, *manners* and *rationality* (see *Table 17*). Three items were used twice for different facets.

Table 17: Scales with item contents for facet and domain scales of the YOUME-List

Names of facet scales	Abbr.	Associated research construct	Item contents	Nb.
Physical attractiveness	PAT	Physical attr*activeness*, vitality	Handsomeness and attractiveness, good body, charisma, personal hygiene, sportiness	5
Sexuality	SEX	Sexual drive, sexual activity, Eros	Lover qualities, sexual responsiveness, carnality	3
Social-economic status	SES	Achievement Striving, social-economic status	Job-related success, financial assets, achievement-striving	3
Responsibility	RES (C)	Dutifulness	Dutifulness, sense of responsibility, reliability	3
Manners	MAN (A)	Order	Decorum, politeness, orderliness	3
Domestic skills	DOS	Cleanliness	Cooking abilities, cleanliness and housekeeper-qualities	2
Gregariousness	GAC (E)	Positive emotions, gregariousness, sociability	Cheerfulness and temperament, gregariousness, humor, charm, *communicativeness*	5
Social adaptiveness	SOA (A)	Compliance, modesty, tender-mindedness	Sensibility and understanding, tolerance, *agreeableness* and likability, affability, *Ability to adjust, subordination, patience, respect*	8
Assertiveness	ASS (E)	Assertiveness, dominance	Dogmaticalness and assertiveness, courage, *ability to dominate, subordination, independency and autonomy*	5
Activeness	ACT (E)	Activity, excitement Seeking, activity	Cheerfulness and temperament, *adventurousness, enterprising spirit, spontaneity*	4
Believing	BEL (O)	Religiousness, spirituality	*Religiousness, spirituality*	2
Artistic disposition	ART (O)	*Openness* to fantasy and aesthetics	Fantasy, sense for art and culture, *variety, musicality*	4
Intellect	INT (O)	*Openness* to ideas, intelligence	Intelligence, general education, creativity, *interest for politics, interest for science*	5
Family orientation	FAM (A)	*Family orientation*	Sense of family, child-friendliness, *desire to have children*	3
Social supportiveness	SOS (A)	Altruism, social support, agape	Sensibility and understanding, helpfulness, *solicitousness, ability to appreciate and to affirm*	4

Emotionality	EMO (A)	Empathy, close-ness, intimacy	*Emotional focus, need for closeness, need to cuddle, agreeableness* and likability, empathy	5
Trustworthiness	TW (C)	Commitment, trust, *trustworthi-ness*	Honesty and sincerity, faithfulness, *respect, trust,*	4
Rationality	RAT (C)	Analytic rational thinking style, pragma	Groundedness, sense of reality and rationality, *solidi-ty*	3
Names of domain scales	**Abbr.**		**Item contents**	**Nb.**
Emotional stability	N-	*Emotional stability*	Psychic stability, mental health, self-assurance, phys-ical health, lack of addictions or alcohol problems, *confidence*	6
Conscientiousness	C	*Conscientiousness*	Dutifulness, sense of responsibility, reliability, deco-rum, politeness, orderliness, groundedness, sense of reality, achievement-striving, economy, *rationality, solidity*	12
Agreeableness	A	*Agreeableness*	*Agreeableness* and likability, affability, empathy, helpfulness, solicitousness, *ability to appreciate and to und affirm, respect, tolerance*	8
Extraversion	E	*Extraversion*	Cheerfulness and temperament, dogmaticalness and assertiveness, gregariousness, humor, charm, courage and bravery, *enterprising spirit, spontaneity, adven-turousness, ability to dominate and superordinate, independency and autonomy, communicativeness*	12
Openness	O	*Openness*	Fantasy and creativity, sense for art and culture, intelligence, general education, sensibility and under-standing, *interest for politics, interest for science, variety*	8

Notes. All scales were rationally constructed as well as supported by findings of multiple Principle Component Analyses (n=423). Items within the long but not within in the short version of the YOUME-L are written in italic type. Every written statement in column 4 should reflect the item wording. Nb.: number of items per scale.

Depending on the number of factors, factor analyses often suggested grouping the same items together, for preferred and for own characteristics. Thus, a structurally relatively stable solution of item-scale classification resulted. 17 factors for preferred characteristics explained 63.0% and 19 factors for own characteristics explained 63.2% of the individual answer variance.

Items which remained on the dimensions, when 2x34 items of the 2x74 items lack, were counted for scaling of the YOUME-L with 2x40 items. Therewith, some scales in the shorter YOUME-L version had fewer items than they had in the long one. A scale was not implemented into the short version, when the correlations between the smaller and the larger scale sank below r=0.85, for the *Wish-Scale* or for the *Self-Scale*, or when the scale had half or less items than the longer version, because the scales were supposed to measure the same or at least highly similar content. Thereby, the short form only implies 11 of the 23 scales of the long version. For these, no new Cronbach's Alpha or scale variances were calculated due to the very high inter-scale correlations with the scales from the long version (see *Table 19*, and *Table 20*).

4.3.2. Retest-reliability

Such as for the BB-PI facets (see *Table 12*), the four-year retest-stabilities (Pearson) for the 13 *Wish-Scales* of the 2x40 items-YOUME-L were mostly moderate (n=250):

- Physical attractiveness .61
- Sexuality .45
- Gregariousness .67
- Socio-economic Status .66
- Family orientation .66
- Domestic skills .43
- Responsibility .50
- Manners .58
- *Emotional stability* .49
- *Conscientiousness* .62
- *Agreeableness* .47
- *Extraversion* .56
- *Openness* .48

4.3.3. Descriptive statistics and subsample differences

Descriptive statistics, intern consistencies and results of MANOVAS for subsample differences (gender, nationality, relationship status) are presented in *Table 18*[20].

Internal consistencies were satisfactory, especially given the relatively small number of items per facet scale. The main effects of gender ($F_{(46, 67)}=4.30$, $p<.001$) and nationality ($F_{(46, 67)}=4.50$, $p<.001$) were significant on the shared sample of 116 partners from 58 couples, while their interaction effect was not ($F_{(46, 67)}=1.09$, $p=.371$). The main effect of relationship status also was not significant ($F_{(46, 268)}=1.23$, $p=.162$), neither was its interaction effect with gender ($F=_{(46, 268)}=1.39$, $p=.059$).

---

[20] Appendix III.6. presents the results of the performed MANOVAS in detail.

Table 18: Descriptive statistics, internal consistency and MANOVAS for scales of the YOUME-List and different (sub-)samples

									Wish-Scales							
	All n=410				German partner n=64		Persian partners n=52		MANOVA Nationality		Females N=222		Males N=138		MANOVA Gender	
Scale abbr.	m	SD	CA	Skew	m	SD	m	SD	F	ηp²	m	SD	m	SD	F	ηp²
Physical attractiveness	4.83	.81	.75	-.12	4.72	.77	5.12	.69	8.64**	.072	4.75	.83	4.93	.77	4.62*	.013
Sexuality	5.19	1.01	.83	-.49	4.97	.91	5.23	1.13	1.96	.017	5.11	1.06	5.30	.89	2.81	.008
Activeness	4.44	.97	.72	-.23	4.42	.89	4.38	.87	.072	.001	4.44	1.00	4.36	.93	.10	.000
Gregariousness	5.21	.74	.70	-.42	5.12	.78	5.00	.63	.728	.006	5.23	.74	5.13	.72	1.33	.004
Assertiveness	4.19	.95	.72	-.36	4.24	.87	4.25	.86	.005	.000	4.45	.86	3.76	.97	46.1***	.114
Socio-economic status	3.90	.98	.72	.00	3.64	1.07	4.32	.84	15.6***	.122	4.09	.98	3.63	.92	18.3***	.049
Believing	2.86	1.87	.77	.64	2.69	1.76	4.40	1.49	30.9***	.216	2.85	1.89	2.93	1.85	.17	.000
Artistic disposition	3.99	.96	.65	-.09	4.01	.91	4.14	.96	.531	.005	3.98	.98	3.99	.94	.16	.000
Intellect	4.26	.86	.66	.11	4.34	.88	3.88	.61	10.2**	.083	4.28	.82	4.18	.88	1.54	.004
Family orientation	5.07	1.47	.86	-.92	5.44	1.15	5.15	1.18	1.70	.015	5.15	1.45	4.93	1.42	1.75	.005
Social supportiveness	5.31	.74	.71	-.07	5.28	.80	5.14	.70	1.08	.010	5.48	.68	5.04	.76	32.6***	.083
Social adaptiveness	5.01	.67	.69	-.23	4.99	.67	5.22	.71	3.17	.028	5.07	.66	4.90	.66	5.06*	.014
Emotionality	5.16	.82	.78	-.32	5.06	.82	4.92	.80	.848	.008	5.22	.82	5.02	.79	5.77*	.016
Domestic skills	3.53	1.30	.71	.10	3.70	1.31	4.29	1.37	6.01*	.051	3.39	1.22	3.78	1.43	6.43*	.018
Trustworthiness	6.23	.69	.75	-1.08	6.23	.60	5.94	.69	6.29*	.053	6.34	.63	6.06	.73	14.1***	.038
Rationality	4.87	.92	.67	-.41	5.04	.90	5.21	.77	1.15	.010	4.92	.94	4.80	.87	1.77	.005
Responsibility	5.54	.82	.88	-.31	5.59	.79	5.84	.81	2.94	.026	5.69	.78	5.32	.81	20.7***	.055
Manners	5.60	.86	.67	-.80	5.60	.77	5.79	.81	1.68	.015	5.68	.89	5.47	.77	5.15*	.014
Emotional stability	5.40	.81	.73	-.73	5.49	.69	5.71	.70	2.78	.024	5.50	.75	5.27	.86	6.94**	.019
Conscientiousness	5.11	.70	.86	-.35	5.19	.66	5.41	.63	3.43	.030	5.19	.69	4.99	.68	6.96**	.019
Agreeableness	5.50	.64	.81	-.16	5.45	.67	5.36	.57	.581	.005	5.63	.60	5.29	.65	25.0***	.065
Extraversion	4.65	.70	.82	-.21	4.62	.64	4.56	.59	.264	.002	4.70	.68	4.52	.69	4.17*	.012
Openness	4.36	.77	.74	.07	4.40	.79	4.06	.64	6.29*	.053	4.39	.75	4.29	.75	1.60	.004

	Self-Scales															
Scale	n=410				German partner n=64		Persian partners n=52		MANOVA Nationality		Females N=222		Males N=138		MANOVA Gender	
abbr.	m	SD	CA	Skew	m	SD	m	SD	F	ηp²	m	SD	m	SD	F	ηp²
Physical attractiveness	4.45	.62	.69	-.15	4.41	.54	4.55	.59	1.88	.016	4.47	.62	4.42	.60	.86	.002
Sexuality	4.58	.81	.76	-.58	4.40	.82	4.58	.88	1.42	.013	4.52	.83	4.67	.77	3.65	.010
Activeness	4.36	.81	.75	-.52	4.25	.73	4.27	.91	.018	.000	4.43	.78	4.24	.84	4.29*	.012
Gregariousness	4.68	.65	.70	-.59	4.59	.55	4.46	.69	1.34	.012	4.71	.65	4.63	.64	1.20	.003
Assertiveness	4.37	.75	.75	-.17	4.42	.69	4.25	.77	1.49	.013	4.30	.76	4.49	.73	6.54*	.018
Socio-economic status	4.10	.73	.65	-.34	4.19	.70	3.94	.94	2.74	.024	4.06	.71	4.17	.73	2.53	.007
Believing	3.19	1.44	.75	.22	3.23	1.34	4.24	1.06	19.6***	.149	3.25	1.42	3.21	1.46	.00	.000
Artistic disposition	4.28	.81	.61	-.35	4.09	.83	4.06	.85	.022	.000	4.34	.79	4.16	.85	4.96*	.014
Intellect	4.43	.74	.66	-.24	4.41	.65	3.84	.73	20.7***	.156	4.36	.73	4.55	.73	6.34*	.017
Family orientation	4.82	1.00	.74	-.87	5.05	.83	4.85	1.00	1.39	.012	4.88	.99	4.76	.97	1.33	.004
Social supportiveness	4.92	.70	.73	-.82	4.86	.58	4.35	.88	13.8***	.110	5.02	.67	4.78	.68	10.32**	.028
Social adaptiveness	4.57	.62	.73	-.49	4.50	.58	4.55	.81	.148	.001	4.61	.61	4.48	.59	4.62*	.013
Emotionality	4.83	.72	.79	-.46	4.66	.73	4.55	.70	.679	.006	4.99	.64	4.56	.74	37.2***	.094
Domestic skills	4.27	1.04	.60	-.57	4.27	.96	4.24	1.06	.036	.000	4.40	.98	4.11	1.07	7.97**	.022
Trustworthiness	5.17	.63	.72	-.45	5.24	.48	4.93	.94	5.07*	.043	5.18	.63	5.17	.59	.06	.000
Rationality	4.63	.77	.67	-.41	4.79	.78	4.55	.88	2.42	.021	4.56	.77	4.72	.78	4.29*	.012
Responsibility	5.11	.70	.89	-.94	5.24	.56	4.96	.83	4.78*	.041	5.20	.65	5.00	.71	8.43**	.023
Manners	5.08	.61	.66	-.81	4.99	.50	5.07	.82	.429	.004	5.15	.58	4.98	.62	8.16**	.022
Emotional stability	4.71	.69	.71	-.52	4.83	.62	4.79	.76	.110	.001	4.72	.65	4.73	.71	.16	.000
Conscientiousness	4.82	.54	.83	-.49	4.89	.44	4.77	.65	1.42	.013	4.85	.51	4.79	.56	.98	.003
Agreeableness	4.93	.59	.81	-.92	4.88	.48	4.50	.82	9.86**	.081	5.00	.56	4.82	.59	8.57**	.023
Extraversion	4.47	.58	.82	-.39	4.43	.48	4.30	.63	1.48	.013	4.49	.58	4.43	.57	.59	.002
Openness	4.46	.63	.71	-.23	4.37	.57	3.96	.64	13.2***	.105	4.46	.63	4.46	.62	.00	.000

Notes. _The Persian and the German subsamples were parallelized in their demographics. The male and the females subsample were the halves of 96 dyads._
_*p<.05. **p<.01. ***p<.001._

4.3.4. Inter-scale correlations

Table 19, Table *20* and *Table 21* reveal similar patterns for correlations between own and between *Wish-Scales*[21]. Predominantly, facets associated with the same *Big Five* trait showed moderate correlations with each other (*convergent validity*) and mostly lower correlations with facets associated with other aspects of the *Big Five* (*discriminant validity*): e.g., there were moderate inter-correlations between the *extraversion*-related facets *activeness, assertiveness, gregariousness, physical attractiveness*, and *sexuality* - for own as well as for preferred partner characteristics.

Moreover, no high correlations between the different facet scales occurred (r<.7), i.e., that the scales substantially differed from each other. This result indicated that the scales measure different traits aspects and thereby justified their discrimination. Conspicuously, correlations between facet and domain scales were higher for *Self-Scales* than for *Wish-Scales*.

Table 21 reveals that, for all scales - except for *assertiveness* and *domestic skills* - the higher one scored in an own trait, the more he / she scored in the preference for this trait: e.g., the higher somebody scored in preferred *physical attractiveness,* the higher he/she also scored in their own *physical attractiveness* with a moderate correlation (PAT: r(407)=.48). Convergent and discriminant evaluation supported this pattern: scales for own and preferred traits with different contents correlated less with each other than those with the same contents.

---

[21] Appendix III.7. presents the inter-scale-correlations, also including those of CCs - with the exact p values.

Table 19: Inter-scale-correlations (Pearson) for Self-Scales with each other (n=410)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
	PAT	SEX	ACT	GAC	ASS	SES	BEL	ART	INT	FAM	SOS	SOA	EMO	DOS	TW	RAT	RES	MAN	N-	C	A	E	O
	(E)	(E)	(E)	(E)	(E)			(O)	(O)		(A)	(A)	(A)		(C)	(C)	(C)	(C)					
s./l.	5/5	3/3	1/4	4/5	2/5	3/3	0/2	2/4	3/5	2/3	2/4	4/8	3/5	2/2	2/4	2/3	3/3	3/3	5/6	9/12	5/8	7/13	5/8
r	1	1	*	.94	*	1	*	*	.81	.90	*	*	*	1	*	.93	1	1	.95	.94	.87	.88	.86
1																							
2	**.46**																						
3	**.38**	**.34**																					
4	**.43**	**.38**	**.52**																				
5	**.32**	.18	**.41**	**.36**																			
6	**.44**	.08	.27	.22	**.44**																		
7	.08	.06	.12	.10	.15	.10																	
8	.28	.19	**.41**	.30	.22	.29	.23																
9	.23	.13	.29	.28	.24	**.45**	-.04	**.57**															
10	.30	.11	.19	.28	.24	.23	.21	.05	.02														
11	.23	.26	.26	**.48**	**.37**	.21	.14	.28	.26	.28													
12	**.40**	**.35**	**.31**	**.43**	**.41**	**.39**	.20	.26	.24	.28	**.56**												
13	.19	**.40**	.19	**.41**	.22	.08	.04	.27	.19	.20	**.69**	**.54**											
14	**.38**	.28	.24	.19	.13	.30	.20	.29	.21	.28	.19	**.34**	.18										
15	.23	.28	.18	**.39**	.26	.14	.07	.09	.11	.28	**.57**	**.49**	**.51**	.09									
16	**.35**	.18	.11	.27	**.36**	**.38**	.09	.20	.28	.25	**.41**	**.54**	.30	**.33**	**.39**								
17	.28	.13	.13	.25	**.39**	**.40**	.18	.16	.18	**.33**	**.53**	**.53**	**.37**	.25	**.54**	**.58**							
18	**.51**	.24	.20	**.41**	**.34**	**.37**	.06	.14	.21	**.35**	**.43**	**.51**	**.33**	.20	**.54**	**.47**	**.55**						
19	**.53**	.23	.25	**.34**	**.42**	**.49**	.26	.20	.21	**.33**	**.33**	**.44**	.19	.25	**.41**	**.50**	**.53**	**.46**					
20	**.45**	.21	.20	**.37**	**.48**	**.55**	.17	.24	**.31**	**.39**	**.52**	**.64**	**.37**	**.34**	**.55**	**.82**	**.83**	**.77**	**.61**				
21	.29	**.31**	.28	**.55**	**.37**	.24	.09	.27	.30	.29	**.90**	**.72**	**.68**	.18	**.66**	**.46**	**.55**	**.56**	**.37**	**.60**			
22	**.48**	**.38**	**.85**	**.79**	**.69**	**.39**	.16	**.43**	**.37**	.28	**.42**	**.48**	**.32**	.23	**.32**	.30	.30	**.38**	**.42**	**.42**	**.47**		
23	.26	.19	**.35**	**.34**	.27	**.42**	.01	**.74**	**.93**	.04	**.36**	**.35**	**.34**	.24	.19	.29	.23	.25	.23	**.34**	**.39**	**.44**	

Notes. s./l.: item number 40-item-version/ in the 74-item-version of the YOUME-L. r: Pearson correlation between these versions. *r<.7 for Self- or Wish-Scale.
r>.3 is printed in bold type.
PAT: Physical attractiveness. SEX: Sexuality. ACT: Activeness. GAC: Gregariousness. ASS: Assertiveness. SES: Socio-economic status. BEL: Believing. ART: Artistic disposition. INT: Intellect. FAM: Family orientation. SOS: Social supportiveness. SOA: Social adaptiveness. EMO: Emotionality. DOS: Domestic skills. TW: Trustworthiness. RAT: Rationality. RES: Responsibility. MAN: Manners. N-: Emotional stability. C: Conscientiousness. A: Agreeableness. E: Extraversion. O: Openness.

Table 20: Inter-scale-correlations (Pearson) for Wish-Scales with each other (n=422)

	1 PAT (E)	2 SEX (E)	3 ACT (E)	4 GAC (E)	5 ASS (E)	6 SES	7 BEL	8 ART (O)	9 INT (O)	10 FAM	11 SOS (A)	12 SOA (A)	13 EMO (A)	14 DOS	15 TW (C)	16 RAT (C)	17 RES (C)	18 MAN (C)	19 N-	20 C	21 A	22 E	23 O
s./l.	5/5	3/3	1/4	4/5	2/5	3/3	0/2	2/4	3/5	2/3	2/4	4/8	3/5	2/2	2/4	2/3	3/3	3/3	5/6	9/12	5/8	7/13	5/8
r	1	1	*	.94	*	1	*	*	.82	.95	*	*	*	1	*	.84	1	1	.99	.99	.93	.85	.90
1																							
2	.42																						
3	.47	.33																					
4	.45	.32	.55																				
5	.36	.24	.37	.33																			
6	.35	.18	.21	.13	.41																		
7	.11	.05	.18	.14	.04	-.01																	
8	.25	.25	.34	.32	.19	.19	.18																
9	.21	.20	.19	.20	.33	.43	-.13	.48															
10	.26	.12	.25	.30	.09	.07	.30	.13	-.02														
11	.21	.27	.19	.40	.18	.12	.03	.29	.18	.22													
12	.22	.20	.13	.31	-.03	.00	.16	.21	.03	.26	.50												
13	.20	.28	.22	.32	.04	.03	.11	.32	.03	.33	.65	.47											
14	.18	.10	.13	.11	.14	.06	.14	.14	.10	.20	.12	.07	.08										
15	.26	.31	.22	.35	.20	.18	.15	.21	.15	.33	.53	.55	.41	.14									
16	.17	.21	.00	.10	.32	.36	-.05	.12	.22	.07	.32	.34	.11	.12	.42								
17	.31	.20	.16	.24	.37	.36	.04	.15	.15	.26	.47	.38	.30	.17	.54	.49							
18	.48	.30	.21	.37	.19	.23	.08	.17	.13	.28	.45	.50	.35	.21	.57	.37	.52						
19	.49	.30	.30	.38	.34	.27	.18	.15	.12	.26	.25	.31	.14	.11	.54	.34	.36	.41					
20	.39	.27	.15	.23	.42	.49	.04	.20	.25	.20	.48	.45	.28	.23	.58	.78	.83	.70	.45				
21	.28	.30	.23	.48	.16	.12	.04	.31	.19	.26	.9	.74	.62	.10	.67	.38	.51	.58	.36	.55			
22	.56	.39	.84	.80	.67	.31	.15	.38	.31	.26	.32	.19	.25	.15	.33	.17	.31	.34	.44	.33	.37		
23	.26	.25	.26	.28	.29	.38	-.04	.70	.92	.05	.33	.21	.27	.14	.26	.23	.23	.22	.16	.31	.35	.37	

Notes. s./l.: item number 40-item-version/ in the 74-item-version of the YOUME-L. r: Pearson correlation between these versions. *r<.7 for Self- or Wish-Scale. r>.3 is printed in bold type.

PAT: Physical attractiveness. SEX: Sexuality. ACT: Activeness. GAC: Gregariousness. ASS: Assertiveness. SES: Socio-economic status. BEL: Believing. ART: Artistic disposition. INT: Intellect. FAM: Family orientation. SOS: Social supportiveness. SOA: Social adaptiveness. EMO: Emotionality. DOS: Domestic skills. TW: Trustworthiness. RAT: Rationality. RES: Responsibility. MAN: Manners. N-: Emotional stability. C: Conscientiousness. A: Agreeableness. E: Extraversion. O: Openness.

Table 21: Inter-scale-correlations (Pearson) between Wish-Scales and Self-Scales (n=409)

S \W	PAT (E)	SEX (E)	ACT (E)	GAC (E)	ASS (E)	SES	BEL	ART (O)	INT (O)	FAM	SOS (A)	SOA (A)	EMO (A)	DOS	TW (C)	RAT (C)	RES (C)	MAN (C)	N-	C	A	E	O
PAT	**.48**	.22	.25	.23	.20	.23	.08	.03	.03	.16	.01	.02	.03	.19	.06	.1	.10	.25	.24	.19	.02	.29	.00
SEX	**.32**	**.66**	.19	.25	.13	.03	.02	.07	.07	.07	.13	.08	.18	.08	.11	.02	.01	.13	.15	.05	.14	.25	.07
ACT	**.35**	.25	**.62**	**.36**	.20	.11	.11	.20	.11	.17	.15	.03	.12	.05	.03	-.09	.06	.02	.15	.01	.12	**.52**	.13
GAC	**.33**	.27	**.41**	**.57**	.11	.04	.08	.19	.11	.23	.28	.19	.25	.08	.22	.03	.10	.23	.21	.11	**.31**	**.47**	.14
ASS	.24	.10	**.31**	.23	.20	.10	.11	.12	-.04	.19	.23	.16	.25	.13	.17	.08	.22	.23	.19	.23	.20	**.33**	.04
SES	.30	.04	.17	.12	.13	**.36**	.08	.06	.08	.14	.00	.09	.04	.17	.07	.14	.17	.26	.19	.26	.03	.18	.08
BEL	.06	.04	.13	.10	.01	-.01	**.83**	.09	-.15	.20	-.07	.12	.01	.08	.10	.00	.04	.00	.19	.04	-.04	.10	-.10
ART	.20	.14	.21	.19	.14	.17	.18	**.53**	**.32**	.02	.07	.06	.11	.11	.05	.06	.07	.05	.13	.10	.08	.25	**.41**
INT	.18	.05	.12	.17	.19	.29	-.08	.28	**.58**	-.03	.13	.01	.00	.11	.08	.14	.09	.09	.13	.15	.14	.21	**.54**
FAM	.26	.04	.23	.23	.07	.11	.23	.03	-.07	**.84**	.11	.16	.20	.12	.27	.03	.22	.23	.25	.17	.15	.21	-.04
SOS	.21	.19	.14	.27	.07	.10	.11	.21	.08	.23	**.50**	.30	**.50**	.12	**.33**	.22	**.33**	.32	.12	**.33**	**.47**	.21	.19
SOA	.25	.25	.21	.24	.12	.11	.15	.07	.00	.16	.23	**.32**	.30	.13	.29	.21	.20	.27	.20	.28	.30	.24	.06
EMO	.14	**.32**	.09	.24	.06	.07	-.01	.18	.07	.19	**.43**	.23	**.60**	.08	.28	.16	.19	.24	.08	.20	**.39**	.17	.18
DOS	.17	.17	.06	.04	.07	.19	.14	.02	.04	.16	-.09	-.03	-.08	.05	.00	.06	.05	.04	.09	.08	-.08	.08	-.01
TW	.22	.26	.17	.25	.15	.12	.07	.10	.05	.21	**.40**	.30	**.39**	.15	**.50**	.23	**.34**	**.43**	.25	**.40**	**.43**	.24	.11
RAT	.21	.15	.03	.12	.14	.14	.01	-.01	.04	.17	.12	.21	.09	.13	.28	**.45**	.29	**.32**	.25	**.44**	.19	.12	.03
RES	.26	.11	.12	.19	.20	.22	.12	.10	.01	.22	.23	.23	.24	.12	**.37**	.3	**.49**	**.42**	.26	**.48**	.26	.21	.07
MAN	**.32**	.11	.18	.24	.11	.19	.04	.02	.00	.22	.26	.27	.23	.14	.30	.23	**.31**	**.59**	.23	**.42**	**.33**	.22	.02
ES	**.31**	.14	.17	.21	.21	.17	.18	.04	-.03	.21	.12	.17	.13	.08	.28	.19	.26	**.31**	**.45**	**.32**	.17	.24	.00
C	**.32**	.15	.14	.22	.20	.24	.10	.05	.03	.25	.24	.29	.20	.17	**.38**	**.40**	**.44**	**.52**	**.31**	**.55**	.30	.23	.06
A	.25	.22	.17	**.31**	.08	.10	.08	.17	.09	.20	**.46**	**.36**	**.48**	.12	**.37**	.25	**.31**	**.38**	.18	**.35**	**.51**	.24	.19
E	**.39**	.27	**.57**	**.49**	.24	.12	.13	.23	.11	.22	.26	.15	.22	.09	.17	.00	.14	.18	.24	.13	.26	**.58**	.15
O	.20	.10	.15	.20	.19	.29	-.04	**.38**	**.55**	-.02	.17	.06	.09	.11	.13	.14	.12	.12	.13	.17	.18	.25	**.58**

Notes. r>.3 is printed in bold type.
PAT: Physical attractiveness. SEX: Sexuality. ACT: Activeness. GAC: Gregariousness. ASS: Assertiveness. SES: Socio-economic status. BEL: Believing. ART: Artistic disposition. INT: Intellect. FAM: Family orientation. SOS: Social supportiveness. SOA: Social adaptiveness. EMO: Emotionality. DOS: Domestic skills. TW: Trustworthiness. RAT: Rationality. RES: Responsibility. MAN: Manners. N-: Emotional stability. C: Conscientiousness. A: Agreeableness. E: Extraversion. O: Openness.

4.3.5. Predictive scale validity

Table 22 presents the correlations (Pearson) of actor and partner variables, fittings (by CCs and divisions) and similarities (by CCs and distances) with *RQ overall*[22]. An example for every part of the table is explained in the following:

- Preferred *physical attractiveness* (*Actor variable, Wish-Scale, PAT*) did not significantly correlate with *RQ overall* for females (r(94)=-.01, p=.889) or males (r(94)=.11, p=.289).

- Own *physical* attr*activeness* (*Actor variable, Self-Scale, PAT*) significantly correlated with *RQ overall*, for females (r(94)=.28, p=.005) as well as for males (r(94)=.42, p<.001 for males).

- Dissimilarity in own *believing* (*Partner distances, Self-Scale, BEL*), negatively correlated with *RQ overall*, for Persians (r(50)=-.55, p<.001) but not for Germans (r(138)= -.04, p=.611). The same accounted for dissimilarity in preferred *believing* (*Partner distances, Wish-Scale, BEL),* which significantly correlated for Persians (r(50)=-.32, p=.022) but not for Germans (r(138)=-.05, p=.620).

- The partners non-fitting in *emotional stability* (*Fitting by division, P1 Wish-Scale, P2 Self-Scales, ES*) negatively correlated with *RQ overall* in the whole sample (r(190)=-.27, p=.008). The actors non-fitting in *emotional stability* (*Fitting by division, P2 Wish-Scale, P1 Self-Scale, ES*) did as well (r(190)=.51, p<.001).

- Regarding own *sexuality* within the whole sample, constellations in which both partners score low (CC ll: r(190)=-.21, p=.004) or where both score high in *Self-Scales* (CC hh r(190)=.32, p<.001) correlated with *RQ overall* the strongest and worked in opposition, while dissimilar constellations did not correlate with RQ significantly (CC_hl_ P1P2: r(190)=.-.14, p=.056; CC_lh_P1P2 low: r(190)=-.05, p=.471).

---

[22] Appendix III.8. presents the correlations to *RQ overall* with the exact p values.

The following main patterns were observed for common scales:

- Many *Wish- and Self-Scales* of both partners correlated positively with RQ, *Wish-Scales* less than *Self-Scales* and actor scales more than partner scales.

- Many invers Similarity-Scales (Distances) and invers Fitting Scales (Divisions) correlated negatively with RQ.

- The listed effects occurred more frequent and stronger for Persians than for Germans, as well as more frequent and stronger for men than for women - but to a smaller extent.

Overall, CCs showed similar predictive validities as the common scales and they were also stronger, as well as more frequently, correlated with RQ for Persians than for Germans and more for men than for women. The following main patterns were observed for CCs:

- *Regarding Similarities*
  - As in *Study I*, the similarity in low scores of socially desirable facets was mainly negatively correlated to RQ (Wish_CC_ll: 10/23, Self_CC_ll: 20/23), while similarity in high scores of the social desirable facets was positively correlated to RQ (Wish_CC_hh: 14/33, Self_CC_hh: 20/23).
  - In comparison, dissimilar values less frequently significantly correlated with RQ (Wish_CC_hl: 7/23, Self_CC_hl: 4/23, Wish_CC_lh: 1/23, Self_CC_lh: 2/23).

- *Regarding Fittings*
  - High *Wish-Scores* with high *Self-Score*s had the highest positive correlations to RQ (*Wish-Self* P1-P2: 18/23, *Wish-Self* P2-P1: 17/23) and high *Wish-Scores* with low *Self-Scores* had the highest negative correlations (*Wish-Self* P1-P2: 15/23, *Wish*-Self P2-P1: 9/23).
  - In comparison, most other combinations of *Wish-Scores* with *Self-Scores* only rarely correlated with RQ.

Table 22: Cross-sectional correlations (Pearson) between scales with the actor's relationship quality for subsamples of partners

| Actor effects: Wish-Scales | | | | | Scale | Partner effects: Wish-Scales | | | | |
All couples n=192	Germans n=140	Persians n=52	Females n=96	Males n=96	name	All couples n=192	Germans n=140	Persians n=52	Females n=96	Males n=96
.05	.08	.02	-.01	.11	PAT	.10	.06	.26	.09	.11
.08	.08	.08	.06	.11	SEX	.17 *	.08	.35 *	.09	.25 *
.02	.01	.05	-.07	.14	ACT	.06	.00	.18	-.01	.12
.21 **	.16	.32 *	.21 *	.21 *	GAC	.18 *	.05	.50 ***	.05	.32 **
.13	.04	.37 **	.16	.11	ASS	.11	.01	.38 **	.03	.20 *
.06	.04	.22	.01	.11	SES	.08	.08	.19	.07	.09
.03	.01	.22	.01	.05	BEL	.11	.08	.36 **	.02	.19
.22 **	.14	.41 **	.15	.31 **	ART	.23 **	.13	.46 **	.07	.38 ***
.19 **	.16	.26	.11	.27 **	INT	.15 *	.12	.20	-.01	.33 **
.17 *	.13	.24	.20 *	.12	FAM	.25 ***	.18 *	.42 **	.07	.42 ***
.18 *	.18 *	.16	.11	.25 *	SOS	.12	.06	.23	.06	.20
.05	.04	.10	.09	.01	SOA	.13	.04	.33 *	.03	.22 *
.14	.18 *	.04	.12	.17	EMO	.12	.04	.27	-.02	.27 **
-.01	.01	.02	.07	-.08	DOS	.12	.07	.30 *	.04	.22 *
.25 **	.22 **	.28 *	.23 *	.28 **	TW	.13	-.03	.40 **	.13	.14
.17 *	.11	.40 **	.21 *	.12	RAT	.22 **	.10	.59 ***	.17	.26 *
.29 ***	.23 **	.46 **	.31 **	.28 **	RES	.20 **	.11	.43 **	.20 *	.21 *
.09	.05	.23	.02	.18	MAN	.10	-.01	.36 **	.16	.04
.19 **	.14	.35 *	.15	.24 *	N-	.16 *	.05	.47 **	.15	.19
.22 **	.13	.50 ***	.22 *	.22 *	C	.22 **	.07	.63 ***	.21 *	.23 *
.18 *	.16	.23	.15	.22 *	A	.13	.04	.36 **	.05	.24 *
.14	.08	.28 *	.10	.18	E	.14	.02	.41 **	.01	.27 **
.20 **	.16	.28 *	.15	.26 *	O	.18 *	.11	.34 *	.02	.36 ***
.22 **	.17 *	.37 **	.16	.29 **	sum *Wish-Scales*	.27 ***	.12	.61 ***	.13	.42 ***
.16 *	.06	.42 **	.29 **	-.04	Imp.friends	.08	-.03	.33 *	-.19	.25 *
.12	.11	.12	.23 *	-.06	Imp. free time	.09	.06	.15	.03	.12
.22 **	.26 **	.21	.21	.24 *	Imp. health	.13	.17	.09	.06	.19
.06	.16	-.10	.06	.06	Imp. material	-.03	.03	-.14	.04	-.09
.10	.12	.07	.08	.13	Imp. job	.02	-.03	.11	-.05	.07
.14	.19 *	.05	.11	.18	Imp. living	.03	.02	.09	.06	.01
.25 **	.23 *	.30 *	.16	.37 **	Imp. family	.22 **	.22 *	.25	.22	.24 *
.25 **	.28 **	.20	.27 **	.22	Imp. partner	.18 *	.15	.25	-.02	.30 **

93

Personality and Relationship Quality

All n=192		Germans n=140		Persians n=52		Females n=96		Males n=96		Scale name	All n=192		Germans n=140		Persians n=52		Females n=96		Males n=96	
Actor effects: *Self-Scales*											Partner effects: *Self-Scales*									
.35	***	.26	**	.58	***	.28	**	.42	***	PAT	.29	***	.15		.64	***	.35	***	.22	*
.30	***	.19	*	.51	***	.25	*	.36	***	SEX	.25	***	.09		.56	***	.23	*	.27	**
.19	**	.02		.51	***	.10		.30	**	ACT	.19	**	.04		.48	***	.20		.19	
.21	**	.15		.28	*	.12		.30	**	GAC	.21	**	.08		.42	**	.20		.23	*
.16	*	.09		.28	*	.12		.21	*	ASS	.07		-.02		.23		.05		.09	
.17	*	.19	*	.12		.06		.30	**	SES	.07		.09		.03		.20		-.05	
.06		-.02		.43	**	.14		-.02		BEL	.08		.02		.37	**	-.05		.21	*
.33	***	.27	**	.47	**	.29	**	.38	***	ART	.27	***	.19	*	.44	**	.17		.37	***
.19	**	.10		.34	*	.04		.36	***	INT	.18	*	.05		.39	**	.13		.23	*
.22	**	.09		.47	***	.19		.26	*	FAM	.24	**	.11		.49	***	.11		.37	***
.26	***	.16		.40	**	.15		.39	***	SOS	.18	*	.10		.26		.22	*	.14	
.29	***	.04		.66	***	.24	*	.35	**	SOA	.29	***	.05		.65	***	.33	**	.25	*
.29	***	.21	*	.47	***	.18		.43	***	EMO	.14	*	.07		.28	*	.18		.12	
.04		-.06		.25		.04		.05		DOS	.05		-.05		.24		-.01		.12	
.47	***	.29	**	.71	***	.49	***	.45	***	TW	.25	***	.06		.48	***	.18		.32	**
.28	***	.17	*	.48	***	.19		.37	***	RAT	.21	**	.09		.43	**	.24	*	.19	
.28	***	.04		.69	***	.25	*	.31	**	RES	.15	*	-.03		.46	**	.17		.14	
.33	***	.13		.66	***	.36	***	.31	**	MAN	.21	**	-.04		.62	***	.21	*	.22	*
.53	***	.41	***	.75	***	.51	***	.55	***	N-	.37	***	.20	*	.68	***	.35	***	.39	***
.32	***	.12		.67	***	.27	**	.38	***	C	.17	*	-.03		.53	***	.20		.15	
.34	***	.16		.58	***	.25	*	.44	***	A	.26	***	.07		.50	***	.27	**	.24	*
.24	**	.09		.49	***	.14		.34	**	E	.21	**	.03		.52	***	.21	*	.22	*
.30	***	.19	*	.51	***	.17		.44	***	O	.24	**	.11		.47	**	.17		.30	**
.45	***	.26	**	.73	***	.36	***	.55	***	sum *Self-Scales*	.33	***	.10		.68	***	.29	**	.38	***
.33	***	.25	**	.54	***	.33	**	.33	**	Satis. friends	.28	***	.23	*	.41	**	.18		.36	***
.31	***	.35	***	.24		.34	**	.25	*	Satis. free time	.33	***	.24	*	.48	**	.18		.42	***
.30	***	.34	***	.30	*	.28	**	.35	**	Satis. health	.25	**	.25	**	.31	*	.31	**	.20	
.36	***	.35	***	.41	**	.38	***	.34	**	Satis. material	.19	*	.13		.38	**	.28	*	.13	
.24	**	.27	**	.19		.08		.49	***	Satis. job	.13		.07		.26		.24	*	.06	
.29	***	.17		.47	**	.18		.44	***	Satis. living	.25	**	.05		.59	***	.28	*	.24	*
.40	***	.23	*	.71	***	.39	***	.41	***	Satis. family	.28	***	.12		.62	***	.41	***	.19	
-.18	*	-.12		-.29	*	-.26	*	-.11		Age	-.19	**	-.15		-.25		-.24	*	-.13	
.03		.08		-.08		-.01		.07		Education level	.03		.03		.02		.04		.02	
-.11		-.16		.03		-.15		-.06		Nb. of children	-.06		-.10		.07		-.03		-.09	
-.03		-.08		.15		-.02		-.05		Household	-.05		-.12		.20		.03		-.14	

All n=192		Germans n=140		Persians n=52		Females n=96		Males n=96		Scale name	All n=192		Germans n=140		Persians n=52		Females n=96		Males n=96	
Wish-Dissimilarities: distances between partners											Self-Dissimilarities: distances between partners									
-.10		.04		-.37	**	-.07		-.10		PAT	-.09		-.01		-.47	***	-.18		-.09	
-.21	*	-.16		-.24		-.19		-.21	*	SEX	-.08		-.14		-.16		-.22	*	-.08	
-.08		-.02		-.14		-.06		-.08		ACT	-.06		-.09		-.10		-.13		-.06	
-.23	*	-.14		-.39	**	-.25	*	-.23	*	GAC	-.07		-.16		.14		-.06		-.07	
-.06		.02		-.21		-.08		-.06		ASS	.02		.00		.01		-.01		.02	
-.10		-.10		-.27	*	-.14		-.10		SES	.16		-.01		.13		-.06		.16	
-.16		-.05		-.32	*	-.05		-.16		BEL	-.14		-.04		-.55	***	-.11		-.14	
-.27	**	-.24	**	-.28	*	-.20		-.27	**	ART	-.25	*	-.19	*	-.31	*	-.19		-.25	*
-.28	**	-.12		-.36	**	-.12		-.28	**	INT	-.02		.09		-.04		.07		-.02	
-.28	**	-.16		-.35	*	-.13		-.28	**	FAM	-.40	***	-.21	*	-.53	***	-.22	*	-.40	***
-.04		-.08		.08		-.02		-.04		SOS	-.23	*	-.12		-.34	*	-.23	*	-.23	*
.08		.02		.04		.01		.08		SOA	-.13		-.04		-.48	***	-.20	*	-.13	
-.19		-.17	*	-.02		-.05		-.19		EMO	-.40	***	-.30	***	-.32	*	-.23	*	-.40	***
-.17		-.18	*	-.05		-.11		-.17		DOS	.09		.17	*	-.11		.07		.09	
-.20	*	-.04		-.48	***	-.22	*	-.20	*	TW	-.32	**	-.17	*	-.59	***	-.42	***	-.32	**
.06		.04		.03		.05		.06		RAT	-.27	**	-.12		-.36	**	-.18		-.27	**
.09		.00		.05		-.04		.09		RES	-.25	*	-.07		-.57	***	-.30	**	-.25	*
.00		.06		-.16		-.04		.00		MAN	-.03		.05		-.52	***	-.14		-.03	
-.33	**	-.15		-.59	***	-.23	*	-.33	**	N-	-.29	**	-.19	*	-.56	***	-.34	**	-.29	**
-.06		-.01		-.28	*	-.08		-.06		C	-.22	*	-.05		-.55	***	-.27	**	-.22	*
.03		-.02		-.05		-.05		.03		A	-.28	**	-.16		-.43	**	-.29	**	-.28	**
-.13		-.03		-.25		-.12		-.13		E	-.01		-.03		-.03		-.08		-.01	
-.25	*	-.14		-.25		-.12		-.25	*	O	-.12		.03		-.21		-.04		-.12	
-.26	***	-.17	*	-.45	**	-.19		-.32	**	sum	-.31	***	-.19	*	-.47	***	-.33	**	-.30	**
.26	***	.17	*	.40	**	.31	**	.21	*	r	.25	***	.26	**	.23		.27	**	.23	*
Non-Fittings: divisions - Wish P1/ Self P2											Non-Fittings: divisions - Wish P2/ Self P1									
-.13		-.11		-.50	***	-.38	***	-.13		PAT	-.33	**	-.21	*	-.42	**	-.26	*	-.33	**
-.12		-.07		-.29	*	-.21	*	-.12		SEX	-.20		-.13		-.27		-.19		-.20	
-.17		-.07		-.46	**	-.29	**	-.17		ACT	-.25	*	-.06		-.46	**	-.19		-.25	*
-.13		-.09		-.19		-.16		-.13		GAC	-.14		-.13		-.16		-.16		-.14	
.01		.04		-.01		-.01		.01		ASS	-.13		-.09		.00		-.02		-.13	
.18		-.02		.10		-.19		.18		SES	-.24	*	-.18	*	.06		.04		-.24	*
-.17		-.03		-.32	*	.00		-.17		BEL	.13		.06		-.22		-.10		.13	
-.32	**	-.14		-.32	*	-.10		-.32	**	ART	-.19		-.21	*	-.36	**	-.33	**	-.19	
-.05		.08		-.23		-.12		-.05		INT	-.21	*	.00		-.23		-.01		-.21	*

All n=192		Germans n=140		Persians n=52		Females n=96		Males n=96		Scale name	All n=192		Germans n=140		Persians n=52		Females n=96		Males n=96	
-.39	***	-.08		-.47	***	-.04		-.39	***	FAM	.00		.02		-.25		-.16		.00	
-.08		-.03		-.23		-.23	*	-.08		SOS	-.39	***	-.13		-.48	***	-.25	*	-.39	***
-.21	*	-.02		-.57	***	-.36	***	-.21	*	SOA	-.27	**	-.03		-.63	***	-.33	**	-.27	**
-.10		-.06		-.24		-.18		-.10		EMO	-.41	***	-.23	**	-.45	**	-.24	*	-.41	***
-.08		.07		-.18		-.01		-.08		DOS	.02		.08		-.06		.03		.02	
-.21	*	.00		-.31	*	-.18		-.21	*	TW	-.40	***	-.21	*	-.60	***	-.44	***	-.40	***
-.13		-.11		-.27		-.24	*	-.13		RAT	-.36	***	-.12		-.36	**	-.10		-.36	***
-.15		.05		-.32	*	-.14		-.15		RES	-.27	**	.01		-.63	***	-.34	**	-.27	**
-.21	*	.03		-.50	***	-.26	*	-.21	*	MAN	-.25	*	-.08		-.58	***	-.39	***	-.25	*
-.27	**	-.10		-.60	***	-.33	**	-.27	**	N-	-.51	***	-.36	***	-.66	***	-.46	***	-.51	***
-.09		.04		-.39	**	-.21	*	-.09		C	-.34	**	-.08		-.60	***	-.27	**	-.34	**
-.18		-.03		-.37	**	-.29	**	-.18		A	-.41	***	-.14		-.58	***	-.35	**	-.41	***
-.16		-.05		-.37	**	-.25	*	-.16		E	-.24	*	-.08		-.37	**	-.20	*	-.24	*
-.17		.04		-.34	*	-.11		-.17		O	-.27	**	-.11		-.38	**	-.20		-.27	**
-.28	***	-.05		-.52	***	-.34	**	-.22	*	sum	-.37	***	-.18	*	-.59	***	-.34	**	-.40	***
.34	***	.33	***	.35	*	.37	***	.31	**	r	.29	***	.29	**	.28	*	.31	**	.28	**
Similarity effects: CC low-low *Wish-Scales*											Similarity effects: CC high-high *Wish-Scales*									
.06		-.09		.38	**	.10		.01		PAT	.08		.03		.21		.01		.15	
-.07		-.09		-.07		-.02		-.13		SEX	.31	***	.21	*	.51	***	.28	**	.34	**
.04		.01		.13		.15		-.08		ACT	.07		-.07		.45	**	-.04		.18	
-.01		.01		-.05		.04		-.07		GAC	.17	*	.07		.40	**	.07		.28	**
-.01		-.01		.02		.05		-.07		ASS	.10		-.11		.47	***	.13		.06	
.07		-.04		.34	*	.10		.04		SES	.13		.18	*	.10		.07		.19	
.02		.03		-.09		.04		.00		BEL	.07		.04		.19		.04		.11	
.00		.11		-.19		.13		-.13		ART	.17	*	.12		.29	*	.06		.29	**
.01		.01		.05		.07		-.06		INT	.19	*	.18	*	.18		.02		.36	***
-.15	*	-.14		-.14		-.11		-.19		FAM	.26	***	.17	*	.50	***	.19		.34	**
-.16	*	-.13		-.19		-.05		-.28	**	SOS	.21	**	.20	*	.22		.13		.29	**
-.21	**	-.17	*	-.31	*	-.14		-.28	**	SOA	.12		.00		.32	*	.12		.12	
-.02		-.05		.04		.02		-.07		EMO	.20	**	.14		.29	*	.10		.29	**
.14	*	.20	*	-.10		.13		.16		DOS	.12		.08		.25		.15		.10	
-.09		-.15		.01		-.03		-.16		TW	.23	**	-.02		.52	***	.23	*	.23	*
-.24	**	-.20	*	-.42	**	-.18		-.31	**	RAT	.12		.02		.42	**	.13		.11	
-.33	***	-.28	**	-.51	***	-.30	**	-.37	***	RES	.26	***	.18	*	.51	***	.27	**	.24	*
-.07		-.08		-.10		-.05		-.09		MAN	.14		.03		.47	***	.14		.13	
-.08		-.10		-.12		-.06		-.11		N-	.17	*	.03		.52	***	.14		.20	*

All n=192		Germans n=140		Persians n=52		Females n=96		Males n=96		Scale name	All n=192		Germans n=140		Persians n=52		Females n=96		Males n=96	
-.25	**	-.18	*	-.45	**	-.20		-.30	**	C	.22	**	.10		.56	***	.23	*	.21	*
-.21	**	-.15		-.33	*	-.09		-.35	**	A	.22	**	.10		.46	**	.15		.28	**
-.01		-.01		-.01		.11		-.14		E	.14	*	-.04		.52	***	.04		.25	*
-.04		-.02		-.06		.07		-.15		O	.20	**	.18	*	.23		.08		.33	**
-.14		-.12		-.21		-.02		-.26	*	sum	.32	***	.16		.60	***	.23	*	.42	***
Dissimilarity effects: CC *Wish-Scales* P1 high – P2 low											Dissimilarity effects: CC *Wish-Scales* P1 low- P2 high									
-.04		.04		-.17		-.04		-.04		PAT	.04		.07		-.01		.03		.06	
-.11		-.03		-.25		-.17		-.06		SEX	.02		.05		-.03		.11		-.10	
-.12		-.09		-.19		-.15		-.12		ACT	-.14		-.12		-.19		-.13		-.17	
-.20	**	-.14		-.30	*	-.21	*	-.19		GAC	-.15	*	-.19	*	-.05		-.12		-.18	
.06		.13		-.16		.03		.09		ASS	.05		.09		-.08		.04		.06	
.01		.02		-.06		-.02		.05		SES	.05		.07		-.03		.00		.10	
-.07		-.01		-.27		-.04		-.13		BEL	.09		.17		-.12		.00		.16	
-.15	*	-.21	*	.02		-.04		-.25	*	ART	-.08		-.18	*	.22		-.12		-.03	
-.11		-.10		-.27		-.07		-.15		INT	-.11		-.16		.00		-.11		-.12	
-.21	**	-.15		-.37	**	-.03		-.42	***	FAM	-.09		-.08		-.14		-.20		.00	
-.07		-.09		.00		-.07		-.06		SOS	-.01		-.05		.10		.01		-.02	
-.12		-.06		-.32	*	-.02		-.21	*	SOA	.02		.02		-.01		-.01		.05	
-.08		.02		-.24		-.01		-.18		EMO	-.04		-.03		-.03		-.06		-.02	
-.16	*	-.13		-.20		-.13		-.19		DOS	.01		.02		.02		-.02		.08	
-.20	**	.06		-.43	**	-.28	**			TW	-.05		.13		-.19				-.07	
-.06		-.05		-.25		-.09		-.03		RAT	.04		.05		-.08		.03		.05	
-.07		-.07		-.08		-.12		.01		RES	-.02		-.04		.01		-.07		.01	
-.09		.06		-.36	**	-.27	**	.09		MAN	.05		.13		-.09		.24	*	-.15	
-.15	*	-.11		-.40	**	-.31	**	.02		N-	-.12		-.12		-.19		-.01		-.24	*
-.13		-.06		-.42	**	-.26	**	.00		C	.00		.00		-.05		.06		-.06	
-.09		-.08		-.10		-.07		-.12		A	.01		.00		.03		-.01		.02	
-.12		-.06		-.23		-.13		-.13		E	-.11		-.11		-.11		-.13		-.09	
-.16	*	-.14		-.28	*	-.15		-.17		O	-.13		-.20	*	.04		-.13		-.14	
-.24	**	-.10		-.57	***	-.25	*	-.23	*	sum	-.04		-.02		-.11		-.01		-.07	
Similarity effects: CC low- low *Self-Scales*											Similarity effects: CC high- high *Self-Scales*									
-.44	***	-.34	***	-.63	***	-.43	***	-.45	***	PAT	.32	***	.18	*	.63	***	.25	*	.40	***
-.21	**	-.11		-.44	**	-.12		-.29	**	SEX	.32	***	.19	*	.58	***	.27	**	.37	***
-.12		-.01		-.39	**	-.06		-.18		ACT	.10		-.03		.44	**	.06		.14	
-.27	***	-.24	**	-.31	*	-.21	*	-.33	**	GAC	.19	**	.16		.28	*	.14		.25	*

All n=192		Germans n=140		Persians n=52		Females n=96		Males n=96		Scale name	All n=192		Germans n=140		Persians n=52		Females n=96		Males n=96	
-.13		-.04		-.26		-.11		-.16		ASS	.07		-.08		.35	*	.01		.12	
-.13		-.01		-.29	*	-.04		-.22	*	SES	.16	*	.12		.23		.14		.17	
-.16	*	-.28	**	.07		-.20		-.12		BEL	.11		.05		.25		.07		.15	
-.19	**	-.17	*	-.21		-.11		-.28	**	ART	.23	**	.15		.43	**	.11		.37	***
-.23	**	-.15		-.32	*	-.16		-.31	**	INT	.11		.04		.36	**	-.02		.25	*
-.16	*	.05		-.61	***	-.11		-.20	*	FAM	.32	***	.20	*	.60	***	.26	*	.38	***
-.21	**	-.22	**	-.17		-.12		-.31	**	SOS	.27	***	.25	**	.38	**	.19		.36	***
-.22	**	-.12		-.44	**	-.18		-.26	*	SOA	.30	***	.18	*	.52	***	.25	*	.36	***
-.23	**	-.12		-.41	**	-.22	*	-.25	*	EMO	.33	***	.29	***	.46	**	.27	**	.40	***
-.09		-.12		-.06		-.16		-.02		DOS	-.04		-.12		.17		-.11		.04	
-.21	**	-.07		-.43	**	-.12		-.30	**	TW	.41	***	.22	*	.72	***	.38	***	.44	***
-.12		-.04		-.29	*	-.13		-.11		RAT	.24	**	.14		.50	***	.18		.29	**
-.13		-.01		-.26		-.12		-.14		RES	.26	***	.08		.58	***	.26	*	.25	*
-.26	***	-.03		-.60	***	-.29	**	-.23	*	MAN	.28	***	.09		.74	***	.26	*	.30	**
-.41	***	-.31	***	-.68	***	-.39	***	-.44	***	N-	.50	***	.39	***	.72	***	.47	***	.53	***
-.16	*	.05		-.48	***	-.18		-.15		C	.31	***	.11		.74	***	.27	**	.34	**
-.26	***	-.15		-.40	**	-.15		-.38	***	A	.34	***	.24	**	.62	***	.26	*	.43	***
-.25	***	-.15		-.42	**	-.18		-.32	**	E	.15	*	.02		.43	**	.10		.21	*
-.29	***	-.28	**	-.29	*	-.20	*	-.38	***	O	.22	**	.14		.41	**	.09		.36	***
-.36	***	-.25	**	-.49	***	-.28	**	-.44	***	sum	.45	***	.27	**	.74	***	.35	**	.56	***
Dissimilarity effects: CC *Self-Scales* P1 high- P2 low											Dissimilarity effects: CC *Self-Scales* P2 low- P1 high									
-.06		-.09		-.01		-.02		-.09		PAT	-.02		.00		-.15		-.13		.13	
-.14		-.27	**	.16		-.15		-.13		SEX	-.05		-.09		.03		-.11		-.01	
.04		.00		.14		.05		.02		ACT	.07		.01		.21		.00		.13	
-.01		-.12		.23		.10		-.10		GAC	-.08		-.12		.02		-.10		-.06	
.01		.01		-.03		-.02		.08		ASS	-.01		.01		-.14		-.06		.02	
-.10		-.14		-.02		-.10		-.11		SES	.05		.02		.10		-.05		.10	
-.09		.08		-.39	**	-.28	**	.00		BEL	-.16	*	-.03		-.37	**	-.09		-.32	**
-.08		-.12		.01		-.03		-.13		ART	-.08		-.13		.04		-.06		-.10	
-.02		-.03		-.01		.08		-.17		INT	-.01		-.05		.06		-.09		.04	
-.21	**	-.19	*	-.26		-.23	*	-.20	*	FAM	-.20	**	-.13		-.36	**	-.11		-.33	**
-.13		-.08		-.18		.04		-.32	**	SOS	-.10		-.08		-.11		-.27	**	.05	
.02		.04		-.06		.06		-.03		SOA	-.09		-.05		-.21		-.15		-.02	
-.17	*	-.20	*	-.10		.09		-.35	***	EMO	-.06		-.06		-.04		-.11		.03	
.00		.04		-.10		.00		.00		DOS	.00		.04		-.11		.03		-.04	
-.19	**	.04		-.46	**	-.28	**	-.07		TW	-.09		.03		-.21		-.03		-.14	

Personality and Relationship Quality

All n=192		Germans n=140		Persians n=52		Females n=96		Males n=96		Scale name	All n=192		Germans n=140		Persians n=52		Females n=96		Males n=96	
-.12		-.04		-.30	*	-.06		-.21	*	RAT	-.12		-.09		-.18		-.23	*	-.04	
-.08		.04		-.35	*	-.14		-.05		RES	-.10		-.04		-.22		-.11		-.07	
.03		.04		.01		-.05		.08		MAN	-.03		-.02		-.08		-.07		.04	
-.17	*	-.18	*	-.17		-.22	*	-.11		N-	-.02		-.01		-.07		-.09		.03	
-.11		-.01		-.37	**	-.11		-.11		C	-.08		-.03		-.23		-.16		.01	
-.13		-.07		-.19		.04		-.33	**	A	-.10		-.03		-.16		-.22	*	.01	
.04		-.01		.21		.09		.00		E	.01		.00		.04		-.07		.10	
-.08		-.10		-.07		.03		-.21	*	O	-.04		-.08		.03		-.12		.03	
-.13		-.09		-.24		-.21	*	-.02		sum	-.18	*	-.16		-.25		-.12		-.23	*

Fittings: CC *Wish-Scales* P1 high - *Self-Scales* P2 high — Fittings: CC *Wish-Scales* P2 high - *Self-Scales* P1 high

All n=192		Germans n=140		Persians n=52		Females n=96		Males n=96		Scale name	All n=192		Germans n=140		Persians n=52		Females n=96		Males n=96	
.16	*	.08		.35	*	.15		.17		PAT	.19	**	.13		.33	*	.11		.27	**
.20	**	.12		.39	**	.12		.27	**	SEX	.27	***	.17		.50	***	.25	*	.28	**
.14		.01		.45	**	.12		.16		ACT	.08		-.10		.50	***	-.02		.21	*
.23	**	.18	*	.34	*	.22	*	.25	*	GAC	.21	**	.16		.31	*	.11		.31	**
.09		-.02		.30	*	.11		.06		ASS	.12		.06		.25		-.06		.27	**
.13		.15		.06		.11		.17		SES	.13		.15		.07		.01		.22	*
.09		.06		.18		.04		.16		BEL	.12		.08		.22		.08		.15	
.22	**	.21	*	.24		.19		.25	*	ART	.14		.09		.24		.01		.27	**
.20	**	.18	*	.26		.15		.27	**	INT	.14	*	.13		.16		.04		.25	*
.23	**	.17	*	.37	**	.22	*	.24	*	FAM	.26	***	.19	*	.44	**	.18		.35	**
.18	*	.17	*	.17		.16		.19		SOS	.15	*	.09		.32	*	.00		.33	**
.25	***	.17	*	.42	**	.30	**	.19		SOA	.26	***	.16		.47	***	.10		.42	***
.16	*	.18	*	.09		.22	*	.10		EMO	.21	**	.15		.35	*	.03		.42	***
-.02		-.06		.08		-.12		.05		DOS	.11		-.02		.39	**	.13		.10	
.30	***	.21	*	.48	***	.20		.41	***	TW	.26	***	.11		.53	***	.25	*	.27	**
.24	**	.22	**	.32	*	.29	**	.19		RAT	.25	***	.17	*	.44	**	.09		.41	***
.30	***	.23	**	.49	***	.30	**	.30	**	RES	.28	***	.15		.59	***	.27	**	.29	**
.24	**	.11		.53	***	.24	*	.23	*	MAN	.20	**	.05		.52	***	.17		.23	*
.33	***	.21	*	.61	***	.33	**	.33	**	N-	.38	***	.31	***	.57	***	.28	**	.48	***
.32	***	.23	**	.53	***	.34	**	.30	**	C	.30	***	.15		.63	***	.22	*	.38	***
.25	***	.19	*	.43	**	.25	*	.26	*	A	.23	**	.13		.51	***	.07		.42	***
.23	**	.10		.50	***	.22	*	.25	*	E	.19	**	.06		.48	***	.05		.34	**
.24	**	.25	**	.22		.25	*	.23	*	O	.21	**	.19	*	.24		.04		.37	***
.37	***	.27	**	.57	***	.34	**	.41	***	sum	.39	***	.24	**	.66	***	.23	*	.55	***

All n=192		Germans n=140		Persians n=52		Females n=96		Males n=96		Scale name	All n=192		Germans n=140		Persians n=52		Females n=96		Males n=96	
Non-Fittings: CC *Wish-Scales* P1 high - *Self-Scales* P2 low												Non-Fittings: CC *Wish-Scales* P2 high - *Self-Scales* P1 low								
-.20	**	-.13		-.35	*	-.43	***	.08		PAT	-.15	*	-.11		-.21		-.06		-.22	*
-.18	*	-.11		-.29	*	-.24	*	-.12		SEX	-.12		-.19	*	-.01		-.15		-.10	
-.26	***	-.17	*	-.44	**	-.34	**	-.21	*	ACT	-.16	*	-.11		-.26		-.14		-.19	
-.24	**	-.20	*	-.29	*	-.34	**	-.12		GAC	-.12		-.18	*	-.02		.08		-.31	**
-.05		-.01		-.09		.01		-.09		ASS	-.03		-.02		-.02		-.10		.09	
-.12		-.06		-.15		-.28	**	.07		SES	-.08		-.05		-.09		-.06		-.10	
-.19	*	-.08		-.34	*	-.12		-.28	**	BEL	-.12		.05		-.38	**	-.31	**	.02	
-.23	**	-.25	**	-.18		-.27	**	-.19		ART	-.18	*	-.20	*	-.12		-.25	*	-.09	
-.11		-.05		-.32	*	-.16		-.08		INT	-.04		-.01		-.12		-.10		.04	
-.19	**	-.13		-.29	*	-.06		-.37	***	FAM	-.14		-.08		-.24		-.20	*	-.09	
-.11		-.07		-.13		-.23	*	.03		SOS	-.07		-.08		-.04		.05		-.20	
-.23	**	-.10		-.40	**	-.21	*	-.25	*	SOA	-.19	**	-.11		-.30	*	-.11		-.26	*
-.05		-.06		-.02		-.03		-.09		EMO	-.10		-.13		-.02		-.01		-.17	
-.13		-.05		-.23		.00		-.25	*	DOS	-.01		.00		.02		-.14		.17	
-.24	**	-.04		-.43	**	-.25	*	-.22	*	TW	-.20	**	.00		-.39	**	-.16		-.24	*
-.14		-.14		-.13		-.23	*	-.08		RAT	-.11		-.08		-.17		-.08		-.16	
-.12		-.03		-.26		-.11		-.13		RES	-.15	*	.02		-.43	**	-.23	*	-.09	
-.18	*	-.06		-.35	*	-.24	*	-.08		MAN	-.06		.08		-.24		-.09		-.05	
-.32	***	-.16		-.56	***	-.44	***	-.17		N-	.34	***	.24	**	.47	***	.20		.46	***
-.21	**	-.11		-.35	*	-.26	*	-.14		C	-.15	*	.00		-.40	**	-.18		-.12	
-.18	*	-.08		-.28	*	-.25	*	-.10		A	-.14		-.10		-.16		-.03		-.24	*
-.27	***	-.15		-.46	**	-.38	***	-.19		E	-.17	*	-.14		-.18		-.09		-.27	**
-.24	**	-.19	*	-.38	**	-.28	**	-.21	*	O	-.16	*	-.14		-.23		-.19		-.13	
-.33	***	-.20	*	-.48	***	-.39	***	-.25	*	sum	-.22	**	-.16		-.28	*	-.18		-.26	*
Fittings: CC *Wish-Scales* P2 low - *Self-Scales* P1 low												Fittings: CC *Wish-Scales* P1 low - *Self-Scales* P2 low								
.06		-.03		.22		.14		-.04		PAT	.06		-.04		.24		.16		-.05	
-.07		-.06		-.11		-.01		-.12		SEX	.01		-.05		.12		.04		-.03	
.07		.09		-.01		.15		-.02		ACT	.17	*	.21	*	.06		.21	*	.13	
.04		.05		.02		.21	*	-.15		GAC	-.04		-.08		.04		-.07		-.01	
-.05		-.17	*	.25		.15		-.24	*	ASS	-.08		-.14		.05		-.23	*	.09	
.04		-.18	*	.50	***	.21	*	-.15		SES	.06		-.07		.36	**	.13		.00	
-.02		-.05		.04		.06		-.10		BEL	.07		.04		.14		.01		.13	
.23	**	.20	*	.29	*	.32	**	.14		ART	.12		.11		.14		.18		.06	
.11		-.04		.47	***	.22	*	-.02		INT	-.01		-.09		.20		-.02		.00	
-.07		-.10		-.02		-.05		-.11		FAM	-.06		-.08		-.02		-.06		-.05	

All n=192	Germans n=140	Persians n=52	Females n=96	Males n=96	Scale name	All n=192	Germans n=140	Persians n=52	Females n=96	Males n=96
.05	.06	.03	.18	-.11	SOS	-.02	-.08	.12	.02	-.07
.06	.03	.11	.19	-.12	SOA	-.01	-.13	.22	-.08	.06
.07	.11	-.04	.20	-.07	EMO	.02	-.05	.21	-.06	.12
.10	.10	.09	.15	.05	DOS	.03	.01	.06	-.05	.10
.02	.02	.02	.02	.01	TW	-.18 *	-.24 **	-.03	-.10	-.25 *
-.03	-.01	-.10	.16	-.23 *	RAT	-.08	-.12	.00	-.12	-.03
-.15 *	-.10	-.34 *	-.19	-.10	RES	-.20 **	-.24 **	-.22	-.19	-.22 *
.00	.00	-.07	.05	-.07	MAN	-.12	-.17 *	-.08	-.15	-.10
-.05	-.05	-.10	.10	-.22 *	N-	-.06	-.07	-.11	-.10	-.02
-.05	-.05	-.13	.07	-.19	C	-.15 *	-.22 **	-.05	-.21 *	-.09
.02	.03	-.03	.13	-.12	A	-.06	-.13	.12	-.02	-.09
.08	.04	.20	.32 **	-.19	E	.05	.02	.11	-.01	.10
.14	.03	.36 *	.33 **	-.07	O	.03	-.04	.19	-.03	.09
.04	-.02	.16	.25 *	-.22 *	sum	-.03	-.13	.19	-.05	-.02
Non-Fittings: CC *Wish-Scales* P1 low - *Self-Scales* P2 high						Non-Fittings: CC *Wish-Scales* P2 low - *Self-Scales* P1 high				
.11	.08	.27	.15	.07	PAT	.07	.07	.04	.01	.14
-.02	-.03	-.01	-.01	-.05	SEX	.01	.05	-.05	.05	-.01
.02	-.08	.21	.10	-.08	ACT	.10	.05	.20	-.01	.20
-.13	-.13	-.12	-.13	-.13	GAC	-.20 **	-.14	-.35 *	-.20 *	-.21 *
.04	.09	-.08	.08	.01	ASS	.01	.01	-.03	-.10	.12
-.09	-.10	-.18	-.03	-.15	SES	-.02	.00	-.19	-.07	.04
.08	.19 *	-.20	-.01	.13	BEL	-.06	.02	-.26	-.02	-.13
-.13	-.11	-.23	-.15	-.10	ART	-.11	-.12	-.10	-.05	-.15
-.03	-.10	.13	-.02	-.07	INT	-.02	-.10	.22	-.02	-.02
-.15 *	-.17	-.16	-.23 *	-.10	FAM	-.21 **	-.19 *	-.33 *	-.08	-.40 ***
.04	.00	.14	.08	.02	SOS	-.02	-.05	.04	-.04	.00
.13	.09	.25	.16	.10	SOA	-.05	-.05	-.10	-.03	-.07
-.14 *	-.18 *	-.06	-.07	-.20	EMO	-.08	-.03	-.18	-.01	-.20 *
.02	-.01	.11	.01	.03	DOS	-.05	-.08	.04	-.05	-.04
.07	.13	-.08		.10	TW	.01	.06	-.12	.01	
-.02	.03	-.28 *	-.04	.01	RAT	-.11	-.06	-.40 **	-.18	-.06
-.05	-.02	-.18	-.05	-.05	RES	-.05	-.04	-.13	-.14	.07
.13	.10	.26	.22 *	.02	MAN	.05	.05	.05	.02	.08
.03	.04	-.04	.14	-.05	N-	.06	.08	-.07	-.03	.19
-.08	-.04	-.31 *	.02	-.21 *	C	-.14	-.08	-.50 ***	-.29 **	-.01
.07	.04	.14	.11	.05	A	-.02	-.05	.04	-.04	.02

All n=192	Germans n=140	Persians n=52	Females n=96	Males n=96	Scale name	All n=192	Germans n=140	Persians n=52		Females n=96	Males n=96
-.03	-.06	.03	.05	-.12	E	-.06	-.07	-.04		-.18	.08
-.07	-.15	.13	-.09	-.04	O	-.03	-.11	.20		.00	-.05
-.07	-.08	-.11	.01	-.14	sum	-.13	-.12	-.32	*	-.18	-.08

Notes. CC: Combination Counts. P1: Actor. P2: Partner of the actor. sum: calculated by adding up corresponding values for all items. Imp.: Importance. Satis.: Satisfaction. Imp. And Satis. Measured by FLZ M: short German version of the questionnaire for life satisfaction (Henrich & Herschbach, 2000).

r: Pearson correlations between different profiles over all corresponding item values.

PAT: Physical attractiveness. SEX: Sexuality. ACT: Activeness. GAC: Gregariousness. ASS: Assertiveness. SES: Socio-economic status. BEL: Believing. ART: Artistic disposition. INT: Intellect. FAM: Family orientation. SOS: Social supportiveness. SOA: Social adaptiveness. EMO: Emotionality. DOS: Domestic skills. TW: Trustworthiness. RAT: Rationality. RES: Responsibility. MAN: Manners. N-: Emotional stability. C: Conscientiousness. A: Agreeableness. E: Extraversion. O: Openness.

*p<.05. **p<.01. ***p<.001.

4.4. Discussion

4.4.1. Conclusion

Overall, the current study has achieved a parallel scaling according to preferences and own characteristics with satisfactory psychometric values for all evaluated parameters:

- Reliability (moderate internal consistency and retest-reliability),
- Intern validity (mostly high discriminant and convergent validity),
- Predictive validity (slight to high predictive scale validities for men, women, Persian, and Germans, for common scales and for the novel CCs).

Test objectivity is given through the standardized conduction and interpretation by a fixed question-answer-scheme and scale classification. Additionally, the moderate correlations between *Wish-Scales* and *Self-Scales* measuring the same content were congruent with the research about mate attraction by similarity (Montoya, Horton, & Kirchner, 2008).

Gender differences in the scales are congruent with previous research findings from major studies: that women prefer higher *socio-economic status,* as well as *assertiveness,* and men prefer more *physical attractiveness* concerning their partners confirms the findings of the theory of evolution (Buss, 1989; Feingold, 1992); In addition, the present study replicates that women score higher in own *agreeableness* and *emotionality* than men as well as lower in their own *assertiveness*, while lower female scores in own *emotional stability* could not be replicated here (Costa Jr, Terracciano, & McCrae, 2001).

The nationality differences in the scales were also congruent with existing research about cultural value systems: Compared to Germans, Iranians are highly religious and score higher in so-called cultural masculinity (Hofstede, 2001; Minkov, 2007). Thereby, they are more driven by achievement, success, and competition rather than by care for others. Correspondingly, the largest differences ($\eta p^2 > 1.0$) which the present study found were that Persians preferred a higher *socio-economic status* and *believing* concerning their partners and their own *believing* was more pronounced, while Germans showed higher scores in *openness, intellect* and *social supportiveness*. The differences between Germans' and Persians' predictive validities were conspicuously large, and are, amongst other topics, further addressed in *Study IV*.

103

4.4.2. Predictive validity

The correlations to RQ resulted in various patterns for the different facets. This validated the attempt to use a more facet-based approach rather than just a domain-based, approach for optimized predictions: e.g., the facet scale measuring preferred *manners* had none, while the facet scale measuring preferred *responsibility* had relevant actor and partner effects. These effects were even stronger than those of the domain scale, which measures preferred *conscientiousness* and contains the items from both mentioned facet *Wish-Scales*.

4.4.2.1. Common scales

The predominant effect congruency with existing research indicated, that the YOUME-L measures similar RQ-relevant aspects as other applied tests. As in previous research in the field, partner similarities, fittings, as well as socially desirable actor and partner characteristics, positively correlated with RQ:

- **Actor and partner effects**

  For the YOUME-L, own and the partner's *Self-Scales* measuring *emotional stability, conscientiousness, agreeableness, extraversion* resulted in slight to moderate correlations to RQ. These findings are congruent with those of a meta-analysis on the *Big Five* (Malouff, Thorsteinsson, Schutte, Bhullar, & Rooke, 2010) and a large study using representative samples from different cultures (Dyrenforth, Kashy, Donnellan, & Lucas, 2010). In former studies, moderately positive actor effects, as well as partner effects, were found and replicated for secure attachment style (Shaver, 1992), self-esteem (Erol, 2014), and *physical attractiveness* (Meltzer, McNulty, Jackson, & Karney, 2014). The corresponding YOUME-L scales *physical attractiveness* and *trustworthiness* also significantly and positively correlated with RQ (*Table 22*).

- **Similarity effects**

  Existing studies showed positive effects of partner similarity, not only in the *Big Five* (Dyrenforth, Kashy, Donnellan, & Lucas, 2010), but also in values and relationship-related attitudes: A study by Gottmann (2014) replicated, slight similarity effects for coping styles longitudinally and an investigation on a representative sample from the Netherlands found similarity in family traditionalism and compassion had a positive effect on RQ (Keizer & Komter, 2015). The YOUME-L also replicated these findings for corresponding similarity scales concerning own *family orientation*, own *emotionality, emotional stability, agreeableness,* and *conscientiousness*.

- **Fitting effects**

    The YOUME-L profile correlations for fittings were significantly associated with RQ, but slightly lower than the moderate ones found in a former cross-sectional study (r=.3 vs. r=.5) by Zentner (2005). A promising follow-up question is if fittings explain the same or additional variance in RQ beyond actor, partner, or similarity variables. *Study IV* provided first results to answer this question.

4.4.2.2. Combination Counts

The finding of *Study I,* that shared weaknesses/strengths affect RQ negatively/positively was replicated. The effect patterns for fittings showed that own partner characteristics mainly significantly affected RQ when this characteristic was highly preferred by the actor. As in *Study I*, the predictive validities of the CCs were similarly high as those of the actor, partner, or distance scores. However, they partly followed a different pattern than the common scales, in that their correlations to RQ could only partly be explained by the pure actor and partner effects: e.g., dissimilarity in scores of preferred *gregariousness* negatively (CC_hl: $r(190)=-.20$, $p=.006$; CC_lh: $r(190)=-.15$, $p=.039$), while similarity in low scores did not (CC_ll: $r(190)=.-.01$, $p=.867$) and similarity in high scores significantly and positively correlated with RQ (CC_hh: $r(190)=.17$, $p=.018$). Thereby, the CCs contributed to give a more differentiated image, which constellations of the partners values are relevant and, thus, should be considered for further explanatory and predictive purposes in future work.

4.4.3. Limitations and outlook

With the help of the developed profound scaling, more elaborated questions about predictive validity and answer distortions in real dating settings can be examined: the current study thus provided the basis for the two application-oriented *Studies IV and V*.

Two major shortcomings of the present study must be taken into account:

(1) Retest-stabilities and correlations to other personality tests (retest-reliability and external validity) were only tested for some - not for all - scales and only for the short - but not for the long - version of the YOUME-L, thus far. This should be made up later. Stability, especially over time, might be relevant for long-term predictions and should be tested thoroughly in the future – as already accomplished for the BB-PI in *Study II*.

(2) Although the sample size was appropriate to back-up a rational test construction, the smaller sample of couples for testing predictive validities did not meet the requirements, as Alpha-errors accumulate due to a large amount of tested correlations (23x8 common

scales, 23x12 CCs). Therefore, while results give an impression of predictive validity, they do not allow 'safe' conclusions about the relation of single traits to RQ - especially not about differences between the even smaller subsamples. However, the differences between cultures and genders might be a promising topic for future research on larger inter-cultural samples.

5. Study IV: *Cross-sectional Prediction of Relationship Quality with Machine Learning by Personality Traits*

Abstract

To what extent is it possible to use ML to predict the success of a relationship based on the self-assessed personality of both partners? The RQ of 192 subjects (96 couples) was reproducibly predicted in a 10-fold CV based on linear additive models. The present study applied the questionnaire designed in *Study III* and outlined it to be promising for measuring actual matching effects: Models based on interaction variables, which combine values of the actor and the partner (6 to 19%), were superior to models based on only actor or partner effects (< 5%).

Additionally, the findings were compared and related to the results of *Study I:* In both studies the models, which were trained either on the basis of variables of *general* or *relationship-related personality* were not able to elucidate significant variance of the RQ, whereas the model based on both was able to elucidate it (9 and 11%). These cross-sectional models predict RQ much less accurately than the longitudinal one model from *Study I* (37%). The current findings contradict *Study I*'s results for the different trait domains: Models, based on variables of *emotional stability*, relevantly predicted the RQ, while models based on interest- and *extraversion*-related variables did not. Moreover, *Study I's* results were expanded: partner preferences, as well as fittings of these preferences to own traits, did not yet incrementally predict RQ in the cross-section and models performed much better for Persians than for Germans.

Keywords: relationship satisfaction, *machine learning*, *personality traits*, dyadic data analysis, linear regression, *Big Five*, *cross-validation*, personality, mate preferences

5.1. *Additional background*[23]

5.1.1. *Fittings*

Study III found similar predictive validities and different effect patterns for fitting but similarities scores (see *Table 22*). Up to now, an open research question concerns, whether fittings have an incremental validity for predicting RQ beyond actor or similarity effects, as well as for whom they are important: e.g., regarding the typical gender differences in mate selectivity (Buss, 1989), it might be more important for females that males fit their preferences than the other way around.

5.1.2. *Nationality as moderator*

Until now, little research has been conducted on RQ and personality within arranged marriages within the region of Iran (Lev-Wiese & Al-Krenawi, 1999). Aside from the result from *Study III*, indicating that Persians' RQ is more strongly affected by personality effects than the Germans RQ', it is an open question, if and which personality effects are more or less important for their RQ. Arranged relationships may lack some major influences of initial attraction or emotions that freely chosen partnerships are affected by. Thus, it may be that actor, partner, fitting and similarity effects are stronger in arranged relationships, as a good 'forced' choice might protect against the negative effect of not having chosen freely.

5.1.3. *Personality traits as moderators*

Only a scarcity of existing research has focused on the question, whose RQ is especially influenced, positively and negatively, by personality effects, e.g., for whom risky traits are riskier and who becomes happy in their relationships nonetheless.

Joel, Eastwick and Finkel (2017), in a study using CVs and other ML methods on a small dataset, tried to predict initial attraction in a mating context – as a very early aspect of RQ - based on such interaction effects of *personality traits*: it could not be relevantly explained by decision tree models that simulate complex interaction effects between own and preferred partner characteristics. However, until now, the question is left open if these

---

[23] Preceding background sections which are additionally relevant for this study:
1.1.2. Reproducible success of previous prediction models
2.1.1. Actor-, partner- and similarity effects
2.1.2. Trait-specific effect of similarities
2.1.3. Effects of relationship-related and *general personality*
2.1.4. Gender effects
Relationship quality, preferred and own personality

'matching' effects could instead reproducibly predict current RQ in an existing partnership - only two studies with a corresponding focus have been sourced:

(1) During a longitudinal examination across the first two years of parenthood (Kohn, et al., 2012), RQ dropped down for highly anxious individuals when their partners perceived themselves as less social supportive, less adaptive and less in a good mood.

(2) Another small study by Hudson and Fraley (2014) indicated that partner similarity matters more for insecure subjects: attachment orientations moderated the association between similarity in *personality traits* and *relationship satisfaction*.

In both studies, interactions seem to involve a typical risky trait (anxiousness, insecurity) combined with a typical protective trait (e.g., social supportiveness, partner similarity).

5.1.4. YOUME-List compared to the Relationship- and Attachment Personality Inventory

RQ models based on the YOUME-L, applied within the current study) compared to those based on the BB-PI , applied in *Study I*, could differ in their performance, as one of the two measurements might in various regards be more or less valid than the other, e.g.:

- The BB-PI has five to eight items per facet that describe specific situations and behaviors, while the YOUME-L directly lists the names of abstract traits for a self-assessment. The situation-based responses the BB-PI asks for might be indispensable for a higher predictive accuracy.

- Additional contents measured by the BB-PI might be incrementally relevant for predicting RQ: e.g., regarding aspects of *emotional stability*, such as *stress-resistance, sensitiveness to critique, helplessness, depressiveness, self-damaging behavior, robustness,* and *anxiety*. The YOUME-L only represents these aspects through the two items: *mental health* and *psychic stability*.

- Of course, conversely, contents which the YOUME-L but not the BB-PI measures, e.g., preferences and fittings, might also be relevant for an incremental predictive validity.

5.1.5. The present study

5.1.5.1. Question

The present study tackled the prediction of RQ based on both partner's self-assessed *personality traits* using classic methods from the ML literature. It aimed to replicate and ex-

tend the cross-sectional findings from *Study I* with the assistance of the more economic personality test designed in *Study III* (YOUME-L). The additive predictive validity was compared to the one of the BB-PI (and PD-I), while the predictive roles of different trait domains, fitting, moderator, similarity, actor and partner effects were evaluated.

A cross-cultural sample of 192 subjects (96 couples) was tested for own and preferred partner characteristics as well as for their RQ. As in *Study I*, the *elastic net* was performed on different sets of variables for the whole sample in a 10x10-fold CV (see *Figure 3*).
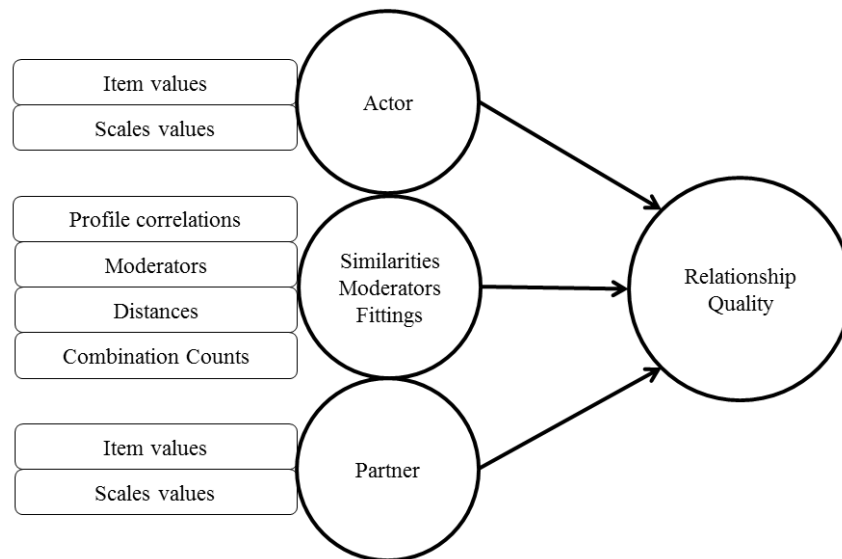
Figure 15. Linear additive model to predict *relationship quality* (right) by different sets of personality variables (left).

In a preceding assessment of predictive scale validity, models predicting *RQ overall* that were based on different scores (e.g., different methods to score partner similarity) or different subsamples (males, females, Persians and Germans) were evaluated.

5.1.5.2. *Hypotheses*

The three following hypotheses for additive linear models were analyzed:

(1) **Reproducible predictive power**

As in *Study I*, ML based models will be able to significantly predict RQ compared to a baseline, which predicts the average of RQ.

(2) **Relationship-related and *general personality***

(2a) As in *Study I*, models based on either variables of general or *relationship-related personality* will not significantly predict RQ compared to the baseline, while models based on both will be able to do so.

110

(2b) Corresponding to existing research (including *Study I)*, models based on *emotional stability*, interests and *extraversion* will significantly predict RQ better than baseline.

(3) **Actor, partner and interaction effects**

(3a) As in *Study I*, actor, partner, similarity and other moderator variables will not be able to demonstrate either an own nor an incremental effect in predicting RQ beyond one another.

(3b) In an additional explorative analysis, the own and additional role of preference and fitting variables for predicting RQ will be evaluated.

The 103 hypothesis-based other moderators are not tested individually, but all of them are tested together in one model and included in the analysis of variable importance. Addressing two research gaps, the hypotheses were based on expectable gender- and nationality related effects as well as on plausible interaction effects between different characteristics[24]:

- **Two characteristics as moderator**

  Except for the first hypothesis, which was derived from a previous study (Kohn, et al., 2012), the following hypotheses concerning two moderating characteristics were based on common sense, i.e. no evidence from existing research was found:

  o Good mood, social adaptiveness, *social supportiveness*, partner fitting in social adaptiveness and *social supportiveness* and similarity in *emotionality* will buffer against the negative impact of emotional instability.

  o Social adaptiveness will moderate actor and partner effects, as it leads to a higher environment-person-adaptiveness.

  o The more relevant a partnership is as a life area for somebody, the more partner fittings and partner traits will influence his or her RQ.

  o *Assertiveness* in one and social adaptiveness in the other partner will complement each other well for better RQ.

  o The RQ of individuals preferring *family orientation* will profit from a partner with high domestic skills and social adaptiveness.

- **Gender and nationality as moderator**

  It is assumed that actor, partner, similarity and fitting variables will have different effects on RQ of Persians and Germans, as mate selection functions differently:

---

[24] Appendix VI.1. presents the complete list of hypotheses-based moderators in the columns AYM to BDE.

- o For Persians' RQ, similarity, actor and partner effects regarding *believing* will be more important than for Germans'.

- o Based on the evolutionary female demanding mating strategy, fitting and partner effects will be more dominant for women.

- o *Physical a*ttr*activeness*, socio-economic status, intellect, *emotionality*, *rationality,* and sexuality will differ in their effect on RQ between men and women, based on their gender-specific relevance found in previous studies.

5.2. Methods

5.2.1. Operationalization

Personality traits and aspects of RQ were tested in a cross-sectional design. The long 74-item version of the YOUME-L was used (see *4.3.1 Scaling*). There were no missing values for personality variables. RQ was tested by the same items described in *Study I* (see 2.2.1.3 *Time two relationship quality*) and calculated as in *Study III (see 4.2.1.3 Relationship quality)*.

5.2.2. Sample

The sample consisted of 96 heterosexual couples, mostly adults with above-average educational levels and living in at least mid-term relationships of at least eight month duration[25]. 26 couples were Persian (27.1%) and 70 couples were German (72.9%). 90 participants (46.9%) had a university degree, 62 (32.3%) had a high-school diploma (German: *Abitur*), and 21 (10.9%) had a lower or no set of qualifications. 19 participants (9.9%) did not state their level of education. 96 participants (50.0%) lived with only their partner, 55 (28.6%) with their partner and at least one child, 14 (7.3%) lived alone, 20 (10.4%) in a shared flat, 4 (2.1%) with their parents and 3 (1.6%) did not state details about their household.

Table 23 shows descriptive statistics concerning relationship-related demographics. RQ measures were positively correlated between partners of a couple (.19 > r < .61): both partners of a couple most agreed about how much they consider a *break-up* (r(82)=.69), while they least agreed about how much they perceived personal constrictions due to one another (r(190)=.19).

---

[25] Appendix IV.2. presents histograms for some of the descriptive statistics.

Table 23: Relationship-related descriptive statistics

r P1-P2	Characteristic	n	Med.	SD	Skew	Error	Min.	Max.
.907***	Age in years	192	3.00	12.1	1.05	.18	16	72
.986***	Relationship duration in months	192	64.5	112	1.95	.18	8	525
.191**	Item: personal constriction (due to the partnership)	192	3.00	.82	-.59	.18	1	4
.574***	Scale of 3 items: *conflicts*	192	4.25	.90	-.70	.18	1	6
.264***	Item: importance partnership/sexuality	159	4.00	.818	-.841	.192	2	5
.487***	Item: satisfaction partnership/sexuality	159	6.00	1.38	-1.21	.19	1	7
.609***	Item: *separation intents*	84	7.00	.813	-1.33	.263	4	7
.509***	Scale on 25 items: partnership *harmony*	84	4.12	.502	-.965	.263	2.55	5
.511***	r between partner's Wish- with others Self-profile	192	.397	.157	-.624	.175	-.12	.65
1.00***	r between partner's profiles (Self-Self, Wish-Wish)	192	.391	.139	-.166	.175	.03	.64

Notes. For scale descriptions and applied questionnaires see Table 4.
*r P1-P2: Intra-couple correlation. r:Pearson correlation. **p<.01. *** p<.001.*

5.2.3. Procedure

The same procedure as in *Study I* was carried out (see *2.2.3 Procedure*)[26]. In addition to similarity scores, fitting scores were calculated and examined: as prepared in *Study III*, fittings combined *Wish-Scales* of one partner with the *Self-Scales* of the other partner - contenting the same trait. Instead of distances, divisions were used for the fittings (see *4.2.1.2 Fittings and similarities*).

5.3.   Results

The internal consistency of the scale for *RQ overall* was excellent (Cronbach's alpha: α=.89, n=84) and, thereby, highly reliable when including the six z-values. The internal consistency of the scale for *RQ overall* with only the three most frequently available z-values was also still acceptably (Cronbach's alpha: α=.69, n=159). Pearson correlations between different RQ measures lay in a moderate-to-high range: this finding emphasized that the measures belong to the same construct (RQ) but still relevantly differ from each other (.41 > r <.62). Descriptive statistics and predictive scale validities with *RQ overall* are presented in *Study III* (see *4.3 Results*)

5.3.1. Score evaluation

To give a first overview and impression, different scores (items, scales, CCs, etc.) and traits (*agreeableness*, *conscientiousness*, *openness*, etc.) were vaguely pre-evaluated in their predictive validity by ML using a rough grid search.

---

[26]Appendix I.5. presents the exemplary Python code.

5.3.1.1. Different traits

Figure 3 presents the predictive performance of the *elastic net* models, based on the actor-, partner, similarity and fitting effects for each separate trait, e.g. the model for *emotional stability* (N-) selected predictors from available common and CC scales for actors N-, partners N-, partner similarity and fitting in N-.

The model based on *emotional stability* relevantly predicted *RQ overall* (N-: MSE=.84), followed by models based on trustworthiness and physical attractiveness (TW/RES: MSE=.90). Models based on other traits did not relevantly manage to predict RQ.

α	λ	trait	MSE
.1	.11	emotional stability	0,84
1	.06	conscientiousness	0,93
1	.07	agreeableness	0,95
1	.05	extraversion	0,95
1	.06	openness	0,95
1	.09	trurstworthiness	0,90
1	.09	physical attractiveness	0,90
1	.01	responsibility	0,92
1	.07	artistic disposition	0,92
1	.05	activeness	0,95
1	.06	rationality	0,95
1	.06	social adaptiveness	0,96
1	.07	emotionality	0,97
1	.06	family orientation	0,97
.55	.06	sexuality	0,97
1	.07	manners	0,97
.55	.05	social supportiveness	0,98
1	.05	intellect	0,98
1	.05	gregariousness	0,99
.55	.04	believing	0,99
1	.04	domestic skills	1,02
.55	.03	socio-economic status	1,02
1	.03	assertiveness	1,04
		Only	0

Figure 16. 10x10-fold *cross-validation* - MSE of the *elastic net* models based on different personality domains predicting the z-value of *relationship quality* overall (n=192)

Notes. 26 Variables in every model: P1 Self-Scale, P1 Wish-Scale, P2 Self-Scale, P2 Wish-Scale, Self-Scale P1-P2 distance, Wish-Scale P1-P2 distance, Fitting P1 Wish scale P2 Self-Scale, Fitting P2 Wish-Scale P1 Self-Scale, all combinations of low and high of Combination Counts for Wish-Wish (4), Self-Self (4), Wish P1-Self P2 (4), Wish P2-Self P1 (4). P1=Partner 1, Actor. P2=Partner 2, Partner.
MSE: mean squared error as deviations between the values which were predicted by the model developed on the train data and the real observed values of the test data; an MSE=1 reflects a deviation of one SD because all variables were z-standardized.

5.3.1.2. Different score types

Figure 17 presents the predictive performance of the *elastic net* models, based on different sets of variables. Models, based on the actors *Self-Scales*, items, all available similarity scores, all available fitting scores, or all variables together, predicted *RQ overall* with equal accuracy (MSE=.82/.83). The model using hypotheses-based moderators predicted *RQ overall* best (MSE=.80), while models based on CCs, partner or Wish-related variables did not relevantly predict RQ.

As actor variables - items compared to scales predicted RQ with a similar accuracy. In contrast to that, for similarities, fittings and hypotheses-based moderators - the different scores predicted RQ more accurately when they were combined in one model: i.e., the different score types for variables, which represent interaction effects, incrementally added up to one another.

For most models using the grid search, the highest tested alpha (α=1) and a low lambda (λ<0.1) were chosen; i.e., that models with high lasso penalty and low ridge penalty better predicted RQ. Therefore, predictive success increased for models which deselected highly correlated variables rather than reduced the number of selected variables by height of their incremental value. To conclude, independency of predictors seemed to be more important than their efficiency.

nb.	α	λ		MSE
16	1.00	.04	Flz	0,97
13	1.00	.04	Demographics	0,99
47	1.00	.10	Main Scores	0,88
1289	1.00	.09	Item-based	0,87
276	1.00	.11	Scale-based	0,85
345	1.00	.10	CC based	0,89
408	.55	.21	Wish-based	0,97
431	1.00	.11	Self-based	0,82
			MAIN SETS	
74	1.00	.09	Items Self P1	0,84
23	.55	.11	Scales Self P1	0,83
74	1.00	.08	Items Self P2	0,93
23	.55	.07	Scales Self P2	0,92
148	.55	.09	Items Self	0,83
46	.55	.11	Scales Self	0,82
			SELF	
23	1.00	.05	Division Scales P1P2	1,01
23	1.00	.08	Division Scales P2P1	0,93
46	1.00	.08	Division Scales All	0,93
74	1.00	.05	Division Items P1P2	1,03
74	1.00	.07	Division Items P2P1	0,92
148	1.00	.07	Division Items All	0,95
23	1.00	.04	Moderators P1P2	0,98
23	1.00	.06	Moderators P2P1	0,95
46	1.00	.06	Moderators All	0,94
242	.10	.29	Fittings P1P2	0,91
219	1.00	.08	Fittings P2P1	0,88
399	.10	.31	Fittings All	0,83
			FITTINGS	
74	1.00	.07	Differences items	1,03
46	1.00	.06	Moderators	1,00
46	1.00	.07	Differences Scales	0,94
439	1.00	.10	Similarities All	0,83
			SIMILARITIES	
			Gender	1,01
25	1.00	.05	Social Adaptiveness	0,88
20	.10	.10	Demogafic	1,01
39	1.00	.05	Nation	1,02
49	1.00	.04	Importance partnership	0,85
22	1.00	.01	Traits	0,84
17	.10	.11	All Hypotheses-Moderators	0,80
123	.55	.11	HYPOTHESES-MODERATORS	
1458	.55	.11	All	0,83
			Baseline	1,01

Figure 17. 10x10-fold *cross-validation* - MSE of the *elastic net* models based on different sets of variables predicting the z-value of *relationship quality* overall (n=192)

Notes. *Nb.: number of variables.P1: Partner 1, Actor. P2: Partner 2, Partner. SOA: social adaptiveness. MSE: mean squared error as deviations between the values which were predicted by the model trained on the train data and the real observed values of the test data. An MSE=1 reflects a deviation of one SD because all variables were z-standardized. Flz: short version of the Life Satisfaction Questionnaire* (Henrich & Herschbach, 2000).

5.3.1.3. *Different subsamples*

Figure 18 presents the predictive performance for the models based on all variables for different subsamples. Persians RQ could be better predicted than the Germans RQ (Persians: n=52, MSE=.73; Germans: n=140, MSE=.89; both with n=46 as training sample)[27]. *RQ overall* was predicted with a similar success for males and females (Males: n=96, MSE=.85; Females: n=96, MSE=.89; both with n=46 as training sample). In conclusion, predictive validity varied between nationalities but not between genders.

n	α	λ			
			Baseline		1,01
52	.1	.57	Germans	0,89	
140	1	.24	Persians	0,73	
96	.10	.35	Females	0,89	
96	.1	.33	Males	0,85	
192	.1	.33	All	0,83	
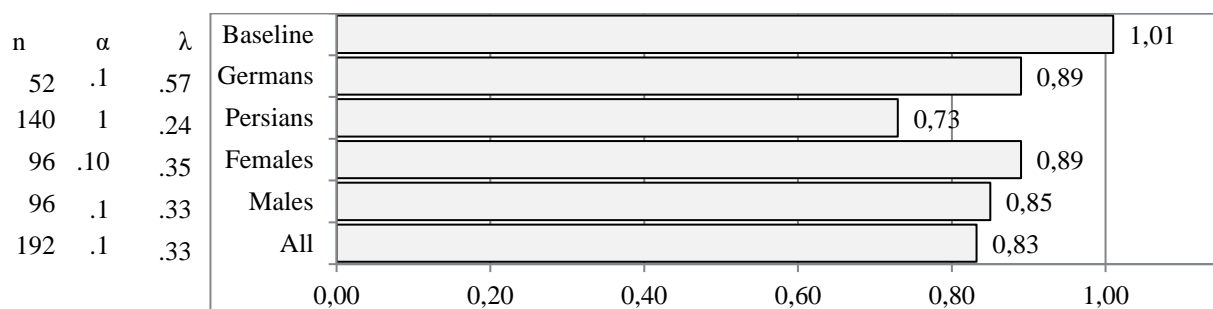
Figure 18. 10x10-fold *cross-validation* - MSE of the *elastic net* models predicting the z-value of *relationship quality* overall for different subsamples

Notes. All 1,335 variables of actor, partner, similarity, and fitting were used. N=46 as a train sample, test sample varying depending on size of subsample. MSE: mean squared error as deviations between the values which were predicted by the model trained on the train data and the real observed values of the test data. An MSE=1 reflects a deviation of one SD because all variables were z-standardized.

5.3.2. *Model performance*

As described in *Study I (see 2.2.3 Procedure)*, a resampled CV set-up was used in combination with an appropriately modified t-test to ensure that the results are reproducible and valid despite the small sample. The reporting of confidence or credibility intervals was omitted because they are not suited to a proper evaluation of results based on repeated CV (Vanwinckelen & Blockeel, 2012). For comparison purposes, a baseline was defined, produced by the performance of a model which always predicted the average of *RQ overall*.

Table 24 presents the predictive performance of models using different combinations of actor, partner, similarity, fitting, and personality domain variables with the more complex parameter tuning[28]. It should be noted that, since model training and model evaluation are carried out on different data sets, r² may become negative.

---

[27] Appendix IV.3 presents the variables which were selected by the models based on Germans and Persians.
[28] Appendix IV.4. presents the model performances in detail.

Table 24: 10x10-fold cross-validation-performance of the elastic net models predicting relationship quality overall

nb.	Predictor variables	Study IV N=192		Study I N=476	
		MSE	r²	MSE	r²
1017	Fittings, similarities, only actor-partner-interactive moderators	.78*	.08		
1439	Actor values, partner values, similarities, fittings, moderators	.85	.02		
1316	Actor values, partner values, similarities, fittings	.91	-.06		
1338	Actor values, partner values, similarities, moderators	.74**	.11	.85	-.03
871	Actor values, partner values, similarities	.77*	.10	.84	-.02
402	Actor values, partner values	.81*	.04	.83	.01
466	Fittings	.81*	.06		
103	Moderators	.67***	.19	.93	-.11
466	Similarities	.77*	.11	.83	.00
201	Actor Values	.81	.04	.83	-.01
201	Partner Values	.92	-.07	.84	-.01
820	Relationship-related and *general personality*	.78*	.09	.74**	.11
589	*Relationship-related personality*	.85	.00	.90	-.08
231	General personality	.86	.00	.76	.07
178	Love	.92	-.06	.86	-.03
355	Values	1.16	-.34	.82	.03
53	Sex	.93	-.09	.87	-.04
268	Interests	1.02	-.22	.62**	.26
117	*Emotional stability*	.76**	.11	.90	-.10
185	*Openness*	.96	-.08	.73	.12
198	*Extraversion*	.86	-.01	.68*	.19
228	*Agreeableness*	.93	-.07	.89	-.08
147	*Conscientiousness*	1.03	-.19	.82	.02
240	Wish	1.37***	-.69		
240	Self	.85	.01		
	Baseline	1.01	-.14	.88	-.05

Notes. Comparison of the results of Study I about the BB-PI / PD-I and Study IV about the YOUME-Lists.
PD-I: Personality Domain Inventory. BB-PI: Relationship- and Attachment- related Personality Inventory.
Different sets of variables were included to evaluate their relevance in predicting relationship quality.
nb.: number of variables the models selected from. r²: forecasting coefficient of determination note that, since model training and model evaluation are carried out on different data sets, r² may become negative.
MSE: mean squared error as deviations between the values which were predicted by the model trained on the train data and the real observed values of the test data; an MSE=1 reflects a deviation of one SD because all variables were z-standardized.
* p<.05, **p<.01, ***p<.001 significantly better than baseline model predicted the average of RQ overall.

(1) Reproducible predictive power:

As hypothesized, some of the models were successfully cross-validated - i.e., predicted RQ better than the baseline. The best model explained 19% of the variance of *RQ overall* (Moderators: MSE=.67, r²=.19; Baseline: MSE=1.01, r²=-.14; p<.001).

(2) Relationship-related and *general personality*:

Appendix IV.5. presents the results of the t-tests, which compare models with one another and against the baseline, in detail.

(2a) As hypothesized, models based on either general or *relationship-related person-ality* variables did not predict RQ significantly better than the baseline (MSE=.85/86, r²=.00, p>.05), while models based on both were able to do so (MSE=78, r²=.09, p=.036).

(2b) As hypothesized, models based on *emotional stability* predicted RQ significantly better than the baseline (MSE=.76, r²=.11, p=.002); however, contrary to the hypothesis, models based on interest- and *extraversion*-related variables were not able to do so (MSE>.85, r²<0, p>.05). As a discriminant validation of the hypothesis, models based on the other personality domains did also not outperform the baseline (MSE>.85, r²<0, p>.05).

(3) Actor, partner and interaction effects:

(3a) As hypothesized, actor, partner, similarity and moderator variables did not demonstrate significant incremental effects in predicting RQ beyond one another. Partner variables (t(99)=0.11 p=.979) did not add any predictive value. Similarity (t(99)=.632, p=.481) and moderator (t(99)=.560, p=.523) variables added predictive value to one another - but not significantly. Supporting the hypothesis, models based on actor and partner variables were not able to predict RQ compared to the baseline (MSE>.80, r²<.05, p>.05); however, contrary to the hypotheses, models based on similarities (MSE=.77, r²=.11, p=.020), fittings (MSE=.81, r²=.06, p=.044) or moderators (MSE=.67, r²=.19, p<.001) were able to do so.

(3b) The predictive performance of the model based on fittings was slightly less accurate than the one from the model based on similarities. Fitting variables did not incrementally contribute to predictive the value of actor, partner and similarity variables: the models even got less accurate by adding these. The model based on Wish-variables performed conspicuously negative (MSE=1.37, r²=-.69, p<.001), while the model based on Self-variables did not predict RQ significantly better than the baseline (MSE=.85, r²=.01, p=.171).

5.3.3. *Variable importance*

To evaluate the importance of different variables, the average weight of each variable was calculated over the 100 models (from the 10x10-fold CV) which were trained using all variables. The most important variables of each variable group are presented in Table 25[29] and are ordered per variable group (actor, partner, interaction effects). To evaluate the importance of the different variable groups, the weights of the variables were added up:

---

[29] Appendix IV.6. presents the whole list of the 600 selected variables with their variable importance. Additionally, the list of variables which were selected by the model based on only the hypotheses-based moderators is presented.

- The weights of actor variables contributed 24.3% to the calculated sum, while 13.7% accounted for own and 7.8% for preferred partner characteristics - this made up a distribution of roughly 2:1.

- The weights of partner variables contributed 16.16% to the calculated sum, while 9.2% accounted for own and 6.0% for preferred partner characteristics - this also made up a distribution of roughly 2:1.

- The weights of similarity variables contributed 59.6% to the calculated sum, while 12.7% accounted for own and 23.46% for preferred partner characteristics - this made up a distribution of roughly 1:1.

- Remaining percentages were due to moderators combining different traits.

Table 25: Variable importance of the top predictors from the elastic net model predicting relationship quality (n=192)

Actor variables (altogether account for 24.31%)				%	β
Self	43	item	protector qualities	1.23	.0429
Self	N	scale	*emotional stability*	1.09	.0379
Self	46	item	sense for art and culture	1.07	.0371
Self	41	item	sexual responsiveness	1.06	.0368
Self	24	item	faithfulness	1.04	.0362
Self	68	item	need to cuddle	.90	.0313
Self	4	item	job-related success	.70	.0242
Self	19	item	sportiness, fitness	.56	.0194
Self	TW	scale	*trustworthiness*	.53	.0185
Self	49	item	need for closeness	.44	.0151
Self	40	item	self-assurance	.32	.0112
Self	3	item	mental health	.22	.0076
Self	28	item	lack of addictions or alcohol problems	.22	.0076
Self	72	item	jealousy	.69	-.0240
Self	25	item	gregariousness	1.81	-.0629
Wish	46	item	sense for art and culture	.80	.0278
Wish	17	item	reliability	.75	.0260
Wish	37	item	down-to-earth attitude	.58	.0200
Wish	48	item	cooking abilities	.22	.0078
Wish	35	item	charm and charisma	.21	.0072
Wish	52	item	subordination	.53	-.0186
Wish	41	item	sexual responsiveness	.51	-.0177
Wish	60	item	respect	.23	-.0080
Wish	67	item	cuteness	1.18	-.0412
Gender * P1 Self INT		moderator	P1 gender * P1 own *intellect*	2.31	.0804

Partner variables (altogether account for 16.16%)				%	β
Self	3	item	mental health	1.34	.0468
Self	7	item	lover's qualities	.70	.0244
Self	55	item	musicality	.63	.0221
Self	4	item	job-related success	.51	.0177
Self	6	item	physical health	.30	.0106
Self	64	item	ability to appreciate and to affirm	.24	.0084
Self	RAT	scale	*rationality*	.22	.0078
Self	25	scale	gregariousness	.28	-.0099
Self	50	item	solicitousness	.31	-.0109
Self	72	item	jealousy	.78	-.0272
Self	30	item	achievement-striving	1.13	-.0392
Self	13	item	honesty	1.33	-.0461
Wish	38	item	economy	1.03	.0359
Wish	47	item	sense of reality and *rationality*	.51	.0179
Wish	50	item	solicitousness	.51	.0178
Wish	42	item	dutifulness	.45	.0158
Wish	66	item	neediness	.43	.0151
Wish	5	item	intelligence	.42	.0146
Wish	31	item	cheerfulness	.21	-.0074
Wish	25	item	gregariousness	.24	-.0082
Gender * P2 Self PAT		moderator	P2 gender * P2 own physical attr*activeness*	.23	.0080
Gender * Sum P2 Wish		moderator	P2 gender * sum of P2 preferences	.31	-.0106
Gender * P2 Self INT		moderator	P2 gender * P2 own *intellect*	.70	-.0242

Similarity variables (altogether account for 59.59%)				%	β
Self	29	item distance	courage	1.38	.0480
Self	50	item distance	solicitousness	1.26	.0439
Self	11	item distance	politeness	.53	.0184
Self	8	item distance	housekeeper qualities	.23	.0081
Self	23	item distance	understanding and compassion	.31	-.0109
Self	70	item distance	desire to have children	1.67	-.0580
Wish	59	item distance	patience	2.42	.0841
Wish	66	item distance	*neediness*	.54	.0187
Wish	50	item distance	solicitousness	.45	.0157
Wish	21	item distance	romantic tendencies	.20	-.0071
Wish	41	item distance	sexual responsiveness	.22	-.0077
Wish	O	scale distance	*openness*	.23	-.0079
Wish	15	item distance	general education	.43	-.0149
Wish	46	item distance	sense for art and culture	.48	-.0167
Wish	19	item distance	sportiness, fitness	1.08	-.0376
Wish	8	item distance	housekeeper qualities	1.36	-.0475
Wish	71	item distance	communicativeness	2.08	-.0722
Self	N	CC high-high	*emotional stability*	2.08	.0722
Self	C	CC low-low	*conscientiousness*	.29	.0100
Self	FAM	CC high-low	*family orientation*	.45	-.0158
Self	N	CC low-low	*emotional stability*	1.19	-.0414
Self	DOS	CC high-high	domestic skills	1.38	-.0479
Self	BEL	CC low-low	*believing*	1.86	-.0647
Self	PAT	CC low-low	physical attr*activeness*	3.66	-.1273

Similarity variables (altogether account for 59.59%)				%	β
Wish	RAT	CC low-low	*rationality*	1.97	-.0685
Wish	MAN	CC low-high	*manners*	1.94	.0673
Wish	BEL	CC low-high	*believing*	1.29	.0449
Wish	DOS	CC low-low	domestic skills	.34	.0117
Wish	N	CC low-low	*emotional stability*	.32	.0112
Wish	INT	CC high-high	*intellect*	.32	.0110
Wish	SES	CC low-low	socio-economic status	.22	.0077
Wish	ART	CC high-low	artistic orientation	.22	-.0076
Wish	DOS	CC high-low	domestic skills	.31	-.0109
Wish	N	CC high-low	*emotional stability*	.36	-.0124
Wish	PAT	CC high-high	physical attr*activeness*	.45	-.0158
Wish	GAC	CC low-high	gregariousness	.46	-.0159
Wish	FAM	CC high-low	*family orientation*	.64	-.0222
Wish	RES	CC low-low	responsibility	.75	-.0261
Wish	DOS	moderator	domestic skills	1.00	.0347
Self	O	moderator	*openness*	.25	-.0089
Self	E	moderator	*extraversion*	1.73	-.0603
Gender * CC P2 Wish high – P1 Self high		moderator	P1 gender * both high preferences	.26	.0090
Gender * Sum Fitting distances		moderator	P1 gender * amount of distances between preferred and own values	.22	-.0076
Nation * CC W high-low/ low/high		moderator	P1 nationality * both combinations of differences in partners preferences	.73	-.0253
Gender * CC P2 Wish low- P1 S low		moderator	P1 gender * combinations of low P2 preferred and P1 own values	2.31	-.0803
P1 Self N * Sim Self EMO		moderator	P1 own *emotional stability* * dissimilarity by distance in own emotionality	3.32	.1153
P1 Wflz partner * r P2 Self- P1 Wish		moderator	P1 importance partnership as life area * profile correlation P2 own with P1 preferred values	2.31	.0803
P1 Wflz partner * CC Fit P1 Wish P2 Self		moderator	P1 importance partnership as life area * combinations of P1 high preferred with P2 low (-) and high (+) values	1.23	.0428
P1 Wish FAM * P2 Self DOS		moderator	P1 preferred *family orientation* * P2 own domestic skills	.33	-.0115
P1 Self SOA * P2 Self ASS		moderator	P1 own social adaptiveness * P2 own assertiveness	.68	-.236

Notes. Model based on all sets of variables (P1, P2, similarities, hypotheses-based moderators) excluding fittings. Overall model: mean squared error=.74, r²=.11. 600 of 984 variables selected.
Table contains variables explaining >.2% of relationship quality on average; sorted by tags of variables as well as by β. β: standardized regression Beta. P1: Actor. P2: Partner. Gender: 0=female, 1=male. Nationality: 0=German, 1=Persian.CC: Combination Counts for all items of one scale. Moderator: product of two z-values.

5.4. Discussion

5.4.1. Conclusion

In this chapter, the major findings are compared with the results of *Study I,* as well as integrated into other preceding research outcomes; replications and substantial differences are systematically analyzed.

5.4.1.1. *Reproducible predictive power*

The most accurate model of the present study (MSE=.67, r²=.19) was slightly less accurate than the most accurate cross-sectional model of *Study I* (MSE=.62, r²=.26). The T1 models of *Study I* and the present study explained less variance than the longitudinal models (max. MSE=.55, r²=.37). This indicated that the cross-sectional, additive predictive validity is larger than the longitudinal one. In line with this finding, a meta-analysis by Malouff et al. (2010), which summarized studies employing simple correlative approaches, showed that the research design (longitudinal or cross-sectional) significantly moderated the effects of *personality traits* on *relationship satisfaction*.

Differences between T1 and T2 did not show in the raw RQ-scale-correlations: For the applied questionnaires - the YOUME-L, the BB-PI and the PD-I – facets with relevant contents similarly and moderately correlated with RQ (see 4.3.5 *Predictive scale validity* and *2.3.1.2. Predictive scale validity*). This finding suggests that the overall predictive quality of matching tests cannot be determined and compared on the basis of separate scale validities but must be evaluated based on all scales to examine how the relevant ingredients add up to one another over time. Models, as they are developed and cross-validated in ML, turned out to be an appropriate approach to validly and reproducibly optimizing the prediction of RQ based on personality data and to assess more accurately the (additive) predictive validity of diagnostic instruments.

5.4.1.2. *Relationship-related and general personality*

The cross-sectional finding from *Study I* that models based either – general or *relationship-related personality* – did not, but models based on both significantly outperformed the baseline, was replicated in the current study. This finding supports the idea that both types of personality are important for RQ beyond one another, i.e., that they may not only explain the same part of the variance, but influence the current RQ in different ways. Contradicting these findings, *general personality* did not contribute to predicting RQ longitudinally in *Study I*. Possibly, the effect of general personality either vanishes over time – at least for the measures of the PD-I – while relationship-related personality becomes more important.

The model based on *emotional stability* was the only trait-based model in the present study which relevantly predicted RQ (MSE=.76, r²=.11). Therefore, the current study and T2 models from *Study I* (MSE=.70, r²=.22) replicated the reoccuring finding (Malouff, Thorsteinsson, Schutte, Bhullar, & Rooke, 2010; Noftle & Shaver, 2006) that *emotional stability* is one of the most important *personality traits* for RQ. That the T1-model based on

emotional stability in *Study I* did not outperform the baseline (MSE=.90, r²=-.10) could be due to the fact that partner preferences are particularly important for cross-sectional prediction of RQ and are measured by the YOUME-L in the current study but not by the BB-PI or the PD-I in *Study I*.

The reason why the model based on interests and *extraversion* for the BB-PI (interests: MSE=.62, r²=.26; *extraversion*: MSE=.68, r²=.19) but not the one for the YOUME-L (MSE>.85, r²<0) may have predicted RQ relevantly, could be that the BB-PI measures many more, concrete hobbies, preferences for activities in free time, and related extraverted and introverted behavior. Nonetheless, the effect did not last over time in *Study I*: the T2-models based on *extraversion* (MSE=.89, r²=0.00) and interests (MSE=.99, r²=-.11) did not relevantly predicted RQ. This finding indicates that their effects on RQ are only temporary and probably not suitable to be included in long-term predictions. Perhaps romantic relationships profit more from similar interests in the short-term, while partners adapt to one another over the long-term anyway. *Study II*'s finding that some interest- and *extraversion*-related facets adjust or develop in the course of relationships supports this idea.

As at T1 in *Study I*, models based on variables related to *conscientiousness*, *openness*, *agreeableness*, values, sex, and love did not or only exceedingly slightly contribute to predict RQ on their own (MSE>.85 , r²<.001). These results contradicted the previous meta-analysis (Malouff, Thorsteinsson, Schutte, Bhullar, & Rooke, 2010), which found that *agreeableness* and *conscientiousness* were associated with RQ in cross-sections, although much slighter than seen with *emotional stability*. A probable explanation is that the (additive) effects concerning these personality domains were too small to be detected in the current small sample.

5.4.1.3. *Actor, partner and interaction effects*

As with the T1-models in *Study I*, actor and partner variables turned out to not - or only exceedingly slightly but not significantly - contribute to predicting RQ and they did not incrementally add up to one another (0-2% of the explained variance of RQ). This finding was inconsistent with existing results on the relevance of actor, partner and similarity effects (Dyrenforth, Kashy, Donnellan, & Lucas, 2010), as well as with the T2 models of *Study I,* which replicated preceding research, that actor effects have the largest and similarity effects have a slight additional relevance for RQ. That these differences in effect sizes did not occur when using a simple correlative approach on the same data in *Study III* indicated that neither the operationalization nor the specific sample was responsible for these results. The separate *personality traits* may already have a visible effect on the RQ already at Time 1, but these

effects seem to add up only in the course of time. When partners are asked why they broke up and they state: "*The little things added up to one another*", they might intuitively refer to this phenomenon. The present study is the first one to support this insight thus far.

Replicating the results of *Study I*, interaction effects explained incremental, but not significantly more variance of RQ beyond actor and partner effects as well as beyond one another (for possible interpretations see *2.4.1.2. Actor, partner, and interaction* effects*).* Nonetheless, in contrast to the findings of the T1-models from *Study I (*MSE>.82, r²<.01*),* all sets of interaction variables - including similarities (MSE=.77, r²=.11), fittings (MSE=.81, r²=.06) and hypotheses-based moderators (MSE=.67, r²=.19) - managed to outperform the baseline. These findings could have one or more of the following reasons:

- The additional interaction variables based on preferences, fittings, and the two different nationalities boosted the predictive performance.

- Persians are more affected by interaction effects than Germans which led to their higher predictive performance for the current dataset.

- The YOUME-L measures more characteristics that are relevant in actor-partner-interaction effects than the BB-PI and the PD-I.

- The scales of the YOUME-L measure characteristics in a way that is more suitable to create predictive interaction effects.

Similar to the distribution from *Study I*, the variable importance for the current models based on all variables - except for fittings - showed much higher weights for actor-partner-interaction (60%) than for actor (24%) and partner (16%) effects. Future work on the top five effects, which each claim more than 2% of the overall variable importance, could be promising. All five are interaction effects:

(1) More so for women than for men, the RQ rises the more they fitted to the low preferences of their partner, by having low own scores.

(2) The more important partnerships were to somebody - as a life area - the better it was for his/her RQ, if the partner fitted his/her preferences.

(3) The more the partner's differed in their preferred *patience*, the higher was their RQ.

(4) Two emotionally, highly stable partners had a higher RQ than partners with other value constellations.

(5) The lower one's own *emotional stability*, the better the similarity in partner's emotional orientation was for the RQ.

The last-mentioned effect corresponded with the results by Hudson and Fraley (2014) who found that insecure subjects profit more from partner similarity.

Preceding research has shown inconsistent results regarding the question of the best scaling for similarities: Occasionally, partner differences turned out to be less strong in RQ-predictions than correlations (Gaunt, 2006); sometimes absolute differences were better predictors (Luo, 2009). Results of *Study I* and S*tudy IV* showed that the different forms of scaling (item and scale differences and moderators) worked best when all put together in one model. Future work should test, if and how different forms of scaling can be combined to further optimize predictive validity with the help of a proper feature selection method.

Last but not least, fittings not only not contributed but even impeded the model performance when actor, partner and similarity variables were considered. Additionally, fitting variables predicted RQ slightly weaker than similarity variables. Therefore, partner fitting might be redundant for cross-sectional predictions of RQ.

5.4.2. Limitations and outlook

In the following sections and in *Table 26*, the limitations and benefits of the present study are juxtaposed and discussed. In summary, future work should contribute to further improvements in long-term predictions of RQ, as well as to increased generalizability in the models developed. As a result, a larger longitudinal (follow-up) study must examine the prognostic validity for the YOUME-L over time with a view to the questions:

- if fittings become additionally predictive later;
- if the effects of *general personality* variables of the YOUME-L will stay important or vanish (as found in *Study I* for the PD-I);
- if the YOUME-L can keep up or even outperform the BB-PI in its predictive validity longitudinally.

Table 26: Study IV evaluation

	Benefits	Limitations
Generalizability	+ Heterogeneous, cross-cultural data was examined. + *Immanent cross-validation of models protects from overfitting.*	- *The sample size was restricted* - *Only partnerships already existing were assessed.* - Cross-sectional design did not enable prediction over time.
Model fit	+ *The elastic net with optimization coefficients alpha and lambda could cope with large amounts of highly correlated variables.* + No missing personality values in the data. + *Dyadic data was handled by allocating partners of one dyad to either train or test sample.*	- *Large number of variables in proportion to the sample size restricts model fit.* - *Only linear effects were analyzed.* - Answers to some questions about *relationship quality* were missing for some partners.
Comparability	+ The number of variables the models selected from and the number they selected were similar. + Sample size was equal in *Study I*.	- Fewer number of personality variables than in *Study I* corresponded with less stable predictions. - Persian subsample might distort the predictions.

Notes. *Evaluation criteria written in italic type also affected Study I.*

5.4.2.1. Sample

The cross-cultural data may have created a more heterogeneous sample and, thereby, reduced the probability of overfitting the models to a specific subgroup of subjects. Nonetheless, due to the large cultural differences in predictive performance, comparability to the results of *Study I* - based on an exclusively German sample - might be restricted. The predictive validity of the models was higher for Persians than for Germans but similar for men and women. This could either mean that (1) the exceedingly small Persian dataset of 52 partners was able to distort the results due to its lack of their representativeness, or that (2) Persian relationships function differently in their relations between personality and RQ, and thus indicate that matching models are transferable cross-gender but not cross-cultural.

Future work should examine with larger samples, to which extent it is possible to develop a cross-culturally transferable model, and which cultural differences reproducibly occur. Moreover, other subsamples, such as risk-couples (e.g., emotionally instable) could be examined to identify, who manages having stable happy relationships nonetheless.

5.4.2.2. Study design

A systematic comparative analysis of models for different subsamples, scores and trait domains was conducted, based on similar number of variables from which the models selected. Due to the intentional economy of the YOUME-L, fewer variables than in *Study I* were used. Since models with higher numbers of predictors typically become more stable, models with less stable accuracies were developed in the present study.

Another relevant limitation of the current study is that only four items of RQ were completed by all attendees. On the one hand, incomplete data yielded a less comprehensive dependent main outcome: As shown in *Study I*, the more items an RQ measure aggregates, the better it can be predicted. Therefore, the missing values could have diminished the model fits. On the other hand, due to the missing values, the RQ was not measured the same way for all participants, and, thus, being restrictedly comparable to the construct of *RQ overall* from *Study I*.

The validity of the findings from *Study I* and from the present study profited from the ML methods in many ways. Mainly, these enabled cross-validated, reproducible results though relatively small sample sizes. When estimating the prognostic validity of personality tests, ML - particularly the advanced mechanisms for predictor selection - could generally contribute to economize them for particular purposes by only choosing only relevant and complementary variables. Predictive relevance – of different contents, numbers of items, wordings of items, different tests and types of scaling – can be more conclusively evaluated, compared, and combined than with common test validation. Additionally, other fields in psychology, which focus on predicting relevant life outcomes or future decisions with the help of *personality traits,* could also profit from working with ML.

6. Study V: *Answer Distortions when Testing Personality Traits in an Online Dating Setting*

Abstract

In the present study, patterns of general and individual *answer biases* in a mating context were tested and described. The short version of the YOUME-L was applied twice within the same sample of 309 participants: once in an anonymous setting (*private*) and once in a faked online dating setting (*public*). High inter-setting correlations for subjects, as well as for all traits, supported a general context-overarching measurement accuracy of self-assessed own and preferred *personality traits*. However, most socially desirable own characteristics - physical attractiveness and *emotional stability* the most - were exaggerated in the *public* setting. Most preferred partner characteristics were not exaggerated, although they were less stable across the settings than own characteristics.

Most subjects (61%) significantly exaggerated answers about their own personality, while only one third answered more modestly in *public* (20%). Around 40% of the subjects significantly exaggerated and around 40% downgraded their partner preferences. Individually differing *answer biases* were found, but an attempt for individual faker-detection based on the answer variances failed. Individual tendencies for self-exaggeration correlated the most with own *emotional stability* and *socio-economic status* (*private*). Individual tendencies for wish-exaggeration correlated the most with preferred *physical attractiveness* in a partner (*private*). *Private* and *public* answers slightly differed in their predictive validity for *relationship satisfaction*: this particularly emphasized the potential practical relevance of *answer biases* for a test application in mating contexts.

Keywords: mating, *personality traits*, online dating, social desirability, impression management, answer distortions, relationship satisfaction

6.1. Background

6.1.1. Validity in high-stake settings

Previous research has shown and discussed the effects of SDR, as well as its possible influences on decision-making in practical psychological settings (Berry, Ones, & Sackett, 2007; Dilchert, Ones, Viswesvaran, & Deller, 2006). *Impression management* refers to a conscious tendency to SDR: own social motives are fulfilled by exaggerating one's positive and understating one's negative characteristics in personally relevant, so-called *high-stake* situations (Dilchert, Ones, Viswesvaran, & Deller, 2006). For instance, aspects of *conscientiousness* are exaggerated, and those of *emotional instability* moderately understated when applying for a job, compared to anonymous settings (Viswesvaran & Ones, 1999).

Meta-analyses indicated that SDR has only a minimal impact on personality tests' criterion-related validity or performance outcomes in the career context, and that it does not depend on individual differences in motivations to manage impressions (Dilchert, Ones, Viswesvaran, & Deller, 2006). One reason for these results could be that it is difficult to affect something that is already very small: the *personality traits* examined to date can only slightly predict job-related performance outcomes, with anonymous as well as with *high-stake* data (Barrick, Mount, & Judge, 2001; Berry, Ones, & Sackett, 2007). While this issue has been extensively studied for job application contexts, it has not been examined thus far, how SDR affects predictive validity of personality tests in *high-stake* mating contexts.

6.1.2.Social desirable responding in online dating

On dating websites, results of questionnaires are typically summarized into a personal profile, which helps potential partners to glean an impression of each other, and which is the basis of partner matching algorithms. An applicative strategy could be to present oneself as maximally appropriate for attracting and being matched with attractive potential partners, based on their suspected preferences. Which characteristics are thereby evaluated as attractive, i.e., socially desirable, could be influenced by suspected widespread partner preferences as they were examined for the YOUME-L in *Study III*: scale averages were pronounced above 'important' for *agreeableness*, *emotional stability*, sexuality, gregariousness, responsibility, *conscientiousness* and manners in a partner (short version of the YOUME-L). Moreover, based on the theory of evolution, it could be influenced by gender-specific preferences for vitality- or resource-related attributes (Rowatt, Cunningham, & Druen, 1999).

Previous empirical studies have documented intentional distortions of self-descriptions in online dating contexts only qualitatively, thus far (Bargh, McKenna, & Fitzsimons, 2002; Brym & Lenton, 2001; Couch & Liamputtong, 2008; Ellison, Hancock, & Toma, 2012; Heino, Ellison, & Gibbs, 2010). Thereby, existing research indicated that online daters exaggerate their self-presentation more often subtly and slightly than non-subtly to enhance their own attractiveness (Bartholomew & Horowitz, 1991). However, a systematic description of quantities and precise pattern of distortions is still lacking and would allow a comparison with analogous work in the field of job application. Moreover it would be a step to check and ensure the transferability of matching results based on anonymous self-assessments to online dating settings.

6.1.3. Individual differences in social desirable responding

Existing research examined different approaches to assess individual tendencies and differences in faking behavior. Among these were several scales, including the *Marlowe-Crowne Social Desirability Scale* (Viswesvaran & Ones, 1999), and the *Balanced Inventory of Desirable Responding* (Leite & Beretvas, 2005). These contain socially desirable item contents that one must disagree with, if one answers honestly and, thereby, socially non-desirably: e.g. '*I never lie*'. In non-applicant samples and reviews, these scores of SDR correlated with the true variance of personality scales: e.g., with the socially desirable traits *emotional stability* and *conscientiousness* (Ones, Viswesvaran, & Reiss, 1996). This finding indicated that they tap more into substance than into response style (McCrae & Costa, 1983). Existing research suggested that the correction of personality and predictor scores with the help of individual answer tendencies, which were assessed by these scores of SDR or by answer variance itself, is rather ineffective (Christiansen, Goffin, Johnston, & Rothstein, 1994; Ellingson, Sackett, & Hough, 1999; Hough, 1998; Ones, Viswesvaran, & Reiss, 1996).

A more promising approach used the covariance index (CVI): Faking is assumed to appear in similar responses to items with social desirable contents in *high-stake* situations; these items would be much less dependent when answered anonymously and honestly. Essentially, fakers are supposed to increase covariance between their responses spuriously and artificially to such items: e.g., an item measuring teamwork might neither be anonymously, empirically nor theoretically associated with an item measuring an aspect of *conscientiousness*. However, these items could strongly correlate with each other in a *high-stake* job condition, as both items refer to desirable work behaviors.

Two studies (Chaney & Christiansen, 2004; Christiansen, Robie, & Bly, 2005) found that the economical CVI (1) could successfully improve criterion-related validity of personality profiles for managerial potential, (2) differed between participants who were instructed to answer honestly and those who were instructed to fake and (3) was sensitive to individual, content- and situation-specific differences in faking. The potential of the CVI for testing faking resistance beyond the job application setting, i.e., for mating contexts, is currently unknown.

6.1.4. Personality correlates

Scores of SDR typically were slightly associated with the *Big Five* (.10<r<.20): the higher somebody scored in *emotional stability*, *extraversion,* and *openness,* and the lower somebody scored in *conscientiousness* and *agreeableness,* the more he/she answered in a socially desirable manner in job application contexts (Furnham, 1986; Ones, Viswesvaran, & Reiss, 1996; Pauls & Stemmler, 2003). Some of the measured traits could have an impact on the degree of faking: e.g. it has been discussed by Dilchert et. al. (2006), that highly conscientious subjects generally perceive more behavioral control in situations and, hence, may be more robust to motivational influences of *high-stake* settings, that increase the willingness to fake. Which *personality traits* are associated with SDR in mating contexts, e.g., in online dating, has not been examined thus far.

6.1.5. The present study

The current study was the first attempt to quantitatively identify general and individual patterns of *answer biases* in a *high-stake* mating context. It was examined how answers about the own personality (*Self-Scales*) and the ideal partner (*Wish-Scales*) differed between an anonymous setting (*private*) and a faked online dating setting (*public*). Thereby, the short YOUME-L was applied twice for 309 subjects. Setting-differences and inter-setting-correlations were analyzed to detect effects of the *high-stake* mating context on responding (e.g., impression management, selection strategies). The three following hypotheses were evaluated:

(1) **Setting-differences**

(1a) In the *public* setting, subjects will exaggerate those own and preferred characteristics, which *Study III*, found to be most attractive and socially desirable in a partner: these will include *agreeableness, emotional stability, manners, sexuality, gregariousness, responsibility,* and *conscientiousness.*

(1b) Concerning *Wish-* and *Self-Scales,* differences and distances between the settings will be similarly high and follow similar patterns, since selective and applicative motives are activated by the mating context: being matched with an attractive partner is probably just as important as being an attractive partner yourself.

(1c) Regarding typical gender differences in mating strategies, men will exaggerate their preferred *physical attractiveness* and own *socio-economic status* more, women will exaggerate their preferred *socio-economic status* and their own *physical attractiveness* more, and no significant discrepancies in setting-differences for other characteristics will occur.

(2) **Correlates of individual answer biases**

(2a) The faking scores - CVI and *public* SDs of answers - will be significantly associated with individual setting-differences and setting-correlations.

(2b) As in preceding research from other *high-stake* contexts, personality facets will significantly correlate SDR: individual setting-differences will be associated with higher *emotional stability, extraversion,* and *openness to experience* as well as with lower *conscientiousness* and *agreeableness*.

(3) **Effects on criterion-related validity**

If *answer biases* influence predictive validity, *private* and *public* scales will differ from each other in how much they correlate with *relationship satisfaction* among individuals living in a partnership – (3a) for *Self-Scales* and (3b) for *Wish-Scales*.

6.2. Methods

6.2.1. Sample

The sample consisted of German adults, most of whom were psychology students. 18 individuals were treated as dropouts due to their extremely low answer variances (SV<.15), more than 75% missing data, or identification in the SPSS outlier analyses. Thereby, the sample was reduced from 327 to 309 subjects. 199 participants were female (64.4%) and 110 were male (35.6%), 84 were single (27.2%) and 225 were committed (72.8%) in relationships which, on average, lasted m=35.9 with SD=51.2 months (Range: 1-520). The average age was m=24.7 with SD=7.29 years (Range: 18 – 65). 44 participants (14.2%) had a university degree, 253 (81.9%) had a high-school diploma (German: *Abitur*), six (1.9%) had finished secondary education, and further six participants (1.9%) did not state their level of education. Matching the norms (Henrich & Herschbach, 2000), most participants stated that *partnership*

and sexuality as a life area was extremely (38.5%), very (42.7%) or quite (13.9%) important to them, while only few (4.9%) rated it as slightly or not important.

6.2.2. Design

Within-subject designs have been evaluated as preferable for researching directed faking (Viswesvaran & Ones, 1999) due to their potential to examine and control for individual differences. Corresponding to this, respondents took the same test twice in the present study. Once, they were instructed to answer honestly in the actual anonymous setting:

> *"Please answer as veridically and honestly as possible, knowing that you give absolutely anonymous particulars. Humans show certain feelings, thoughts, ideas, and behaviors differently strong and frequent. Since subjects, their experiences and life situations are different, there are no true or false answers. Please do not consider first which answer would give the best impression, instead give the answer which most fits to your personally. Most subjects tend to slightly overestimate their selves, which is totally normal. Consider this for your answers and work speedily but carefully."*

Then they were instructed to answer in a fake, non-anonymous online dating setting:

> *"Please, imagine the following very realistically, no matter what you think about such a method to look for a partner and no matter, if you would use it for your own case or not: You answer the following questionnaire within the scope/framework of a serious online dating agency applying a matchmaking algorithm. Thus, on the basis of your NON-ANONYMOUS answers, (1) potential partners that fit to you assessed by the matchmaking algorithm are filtered, and (2) a summary of your personality profile is compiled and visible to other members of the dating website."*

The order of both settings was counter-balanced across participants. They completed the survey within 30 to 40 minutes, while short breaks of approximately ten minutes were allowed before the user was logged out from the survey website.

6.2.3. Operationalization

The scaled 2x40 YOUME-L was used to test own and preferred partner characteristics in a parallel test design (see *Scaling*). In addition to demographics, 'general satisfaction with sexuality and partnership' (*relationship satisfaction*) was measured by one seven-stepped item from the short version of the Life Satisfaction Questionnaire, the *FLZ M* (Henrich & Herschbach, 2000).

6.2.4. Covariance index

To capture individual differences in artificial covariance, the CVI was calculated by totaling the cross-products of the standardized item pair scores. Item pairs were chosen when their correlations were not significant between one another in the *private* setting (p>.1), but in

the *public* setting (p<.01): by this, 150 item pairs were selected[30]. Test-takers were classified as fakers if their CVI was more than two SDs above the average CVI within the sample.

6.2.5. Procedure

6.2.5.1.　　Setting-differences

To analyze mean-level-differences between *private* and *public* scores, an inner-subject-MANOVA was run with the 11 characteristics as dependent variables, the setting as within-subject and gender plus relationship-status as between-subject factors. Profile SD, overall score of answer values, and CVI were also integrated as dependent variables. If a factor had no significant main effect, the MANOVA was run again without this or these factors for the detailed analysis. Every following MANOVA was conducted in the same manner.

6.2.5.2.　　Correlates of individual answer biases

The setting-differences and setting-distances (difference without the sign) between answers of the *private* and the *public* setting were summed up for every person and for every *Wish-Scale* and *Self-Scale* to measure individual *answer biases*. An inner-subject MANOVA was run with all setting-differences and setting-distances as dependent variables: once for *Wish-Scales*, once for *Self-Scales* as within-subject factors and gender plus relationship-status as between-subject factors.

Moreover, individual profile correlations with all item values and correlations for each trait scale over all individuals (Pearson) were calculated. *Public* and *private Wish-Scales* and *Self-Scales,* as well as demographics, were correlated (Pearson) with individual setting-differences, setting-correlations, CVI and SDs of the answers.

Up to this point, the data analysis was based on all 309 subjects. Taking this relatively large sample size and the many tests into account, a strict level of significance of p<.01 was used to focus on effects of a practically relevant size.

6.2.5.3.　　Effects on criterion-related validity

The item value for *relationship satisfaction* was correlated with all *Wish-Scales* and *Self-Scales* (Pearson). Calculations for this last section were only run for the 225 individuals who currently had a partnership. Due to the lower sample size, the less strict level of significance of p<.05 was used.

---

[30] Appendix V.2. presents the selected item pairs and the item correlations to one another in both settings.

6.3. Results

6.3.1. Setting-differences

The inner-subject MANOVA revealed a significant main effect of the setting on the answers ($F(27, 281)=10.99$, $p<.001$), but no interaction effect between setting and gender ($F(27, 281)=1.18$, $p=.253$). The MANOVA did not show any significant differences between individuals who were single and those who lived in a partnership ($F(27, 280)=1.17$, $p=.263$, $\eta p^2=.101$)[31]. Therefore, it was decided to include rather than dismiss data from non-single participants. *Table 27* shows a juxtaposition of the answer shifts between the settings, for *Wish-Score*s (left) and *Self-Score*s (right).

Table 27: Patterns of answer differences between private and public setting (N=309)

								Wish-Scores / Self-Scores							
private		*public*		between *private-public*				*private*		*public*		Between *private - public*			
m	SD	m	SD	p	ηp²	F		m	SD	m	SD	p	ηp²	F	
4.63	.661	4.70	.624	.007**	.023	7.37	PAT	4.55	.682	4.68	.620	<.001***	.119	41.55	
4.92	.777	4.87	.769	.186	.006	1.76	SEX	4.59	.904	4.69	.808	.002**	.032	10.19	
4.86	.606	4.86	.595	.870	<.001	.027	GAC	4.66	.655	4.78	.618	<.001***	.088	29.70	
3.65	.807	3.76	.816	.001**	.034	10.86	SES	4.00	.745	4.10	.734	.003**	.028	8.97	
5.05	.728	5.03	.715	.634	.001	.227	RES	5.01	.790	5.08	.727	.061	.011	3.54	
5.00	.693	5.03	.678	.227	.005	1.46	MAN	5.07	.554	5.15	.547	<.001***	.046	14.96	
4.66	.740	4.71	.732	.098	.009	2.75	N-	4.66	.806	4.87	.734	<.001***	.131	46.47	
4.74	.614	4.73	.601	.526	.001	.402	E	4.59	.634	4.69	.612	<.001***	.069	22.98	
4.43	.721	4.45	.728	.464	.002	.537	O	4.59	.705	4.64	.682	.029	.015	4.83	
5.03	.530	5.04	.496	.967	<.001	.002	A	4.91	.587	5.01	.554	<.001***	.074	24.62	
4.66	.608	4.67	.617	.825	<.001	.049	C	4.77	.550	4.83	.526	<.001***	.040	12.92	
1.13	.254	1.09	.254	<.001***	.063	20.79	SD all³	1.07	.240	1.00	.228	<.001***	.187	70.86	
4.53	.425	4.55	.444	.471	.002	.522	m all³	4.59	.396	4.69	.393	<.001***	.179	66.94	

Notes. *Results of the inner-subject MANOVA (without gender as between subject factor).* df=1, 308.
PAT: Physical Attractiveness. SEX: Sexual Drive. GAC: Gregariousness. SES: Socio-Economic Status.
RES: Responsibility. MAN: Manners. N-: Emotional stability. E: Extraversion. O: Openness. A: Agreeableness.
C: Conscientiousness. Scale values allocated: 4: more true, 5: true,6: very true. ³average over all items. ηp²:
*partial Eta². ** p<.01. *** p<.001.*

As shown in *Table 27*, participants scored significantly higher in the *public* than in *private* setting overall and for nine of 11 *Self-Scales*. These were the exaggerated characteristics ranked by size:

- About ¾ of a SD was noted for own *emotional stability* (N-, d=.777) and *physical attractiveness* (PAT, d=.735).

---

[31] Appendix V.3. presents the results of the MANOVAS in detail.

- Slightly more than ½ SD was noted for *agreeableness* (A, d=.565), *extraversion* (E, d=.545) and the subordinated facet *gregariousness* (GAC, d=.621).

- Slightly more than 1/3 of a SD was noted for *sexuality* (SEX, d=.364), *socio-economic status* (SES, d=.340), *manners* (MAN, d=.439) and *conscientiousness* (C, d=.408).

- With less than 1/3 of an SD, *openness* (O, d=.247) and *responsibility* (RES, d=.210) were not significantly exaggerated.

Only two *Wish-Scales* (see *Table 27*, left side) were significantly exaggerated in *public,* with slightly more than 1/3 of a SD: *socio-economic status* (SES, d=.375) and *physical attractiveness* (PAT, d=.307). The other *Wish-Scales* did not show significant shifts between the settings.

The inter-setting-correlations were high for all Wish-Scales and Self-Scales (range of r(307): 0.632 to 0.845)

6.3.2. Individual answer biases

6.3.2.1. Distribution of individual answer biases

Based on a 99% confidence interval surrounding zero (=+/-.03094) for the individual settings-differences of *Self-Scores*, 189 participants (61.2%) significantly exaggerated while 67 (21.7%) significantly downgraded their answers *public*ly. Only 53 participants (17.1%) did not significantly distort their *Self-Scores* into a clear direction. Based on a 99% confidence interval surrounding zero (=+/-.0416) for the individual settings-differences of *Wish-Scores*, 129 participants (41.7%) significantly exaggerated, while 132 (42.7%) significantly downgraded their answers *publicly*. Only 48 participants (15.6%) did not distort their *Wish-Scores* into a clear direction. *Figure 19* shows the distributions of these individual setting-differences for both *Self-Scores* and *Wish-Scores*.
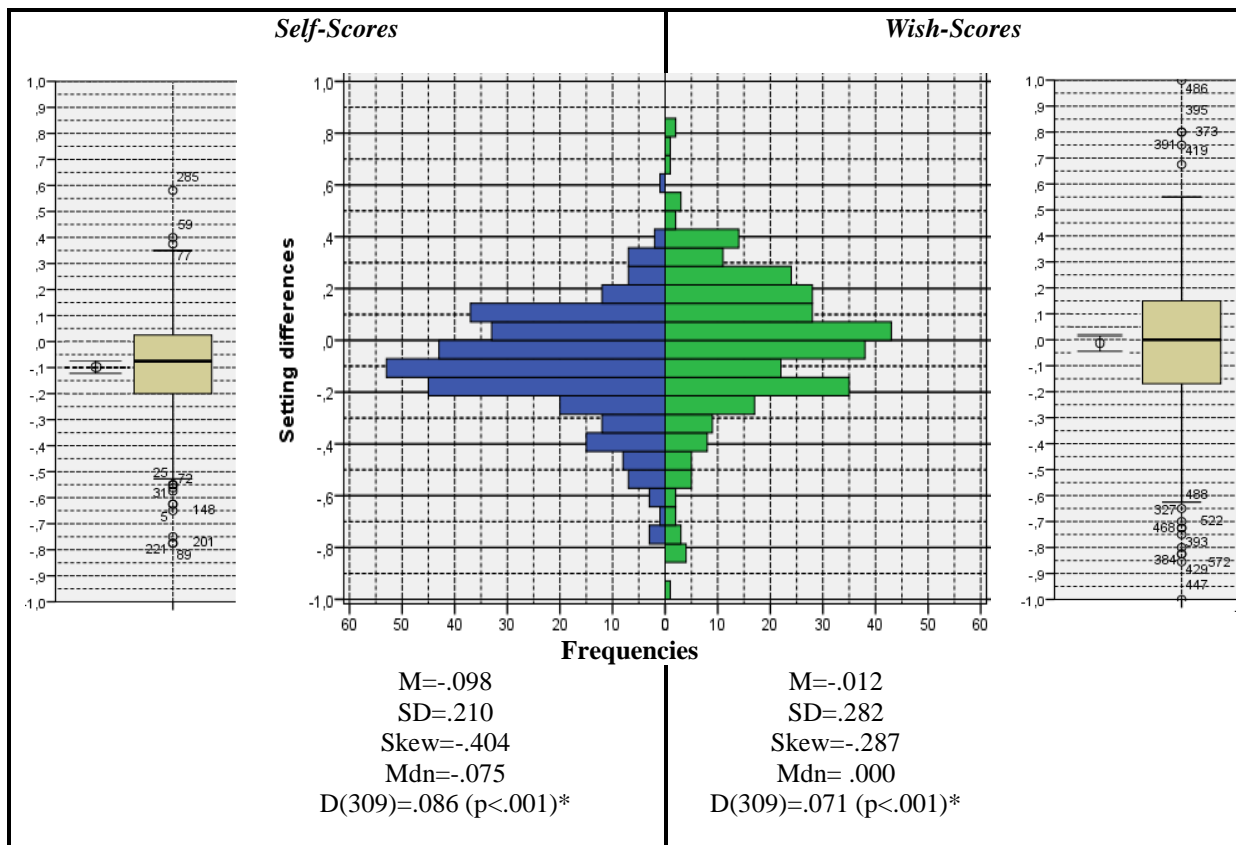
Figure 19. Distributions for individual setting-differences (n=309).

Notes. *For Wish-Scores and Self-Scores, answers between private and public settings were substracted and totaled per person. Histograms (centre), boxplots, averages and confidence interval of 95% (margins) are displayed. *p<.05 in Kolmogorov-Smirnov test for normal distribution.*

Concerning *Wish-Scores*, the individual setting-correlation was high for 107 participants (34.6%: r>.7), medium for 136 (44.0%: .5<r<.7) and low for 66 subjects (21.4%: r<.5). Concerning *Self-Scores*, the individual setting-correlation was high for 180 participants (58.3%), medium for 133 (33.3%), and low for 26 subjects (8.4%). Figure 20 shows the distributions of these individual setting-correlations for both *Self-Scores* and *Wish-Scores*[32]. Individual setting-correlations of *Self-Scores* and *Wish-Scores* correlated moderately with one another (r(307)=.34, p<.001): i.e., the more somebody distorted the answers about preferred characteristics, the more he/she also distorted answers about own characteristics. Yet, subjects who exaggerated *Self-Score*s more did not exaggerate *Wish-Score*s more or less (r(307)<.10, p=.095).

---

[32] Appendix V.4. presents some histograms, further descriptive statistics and inter-setting correlations per characteristic.
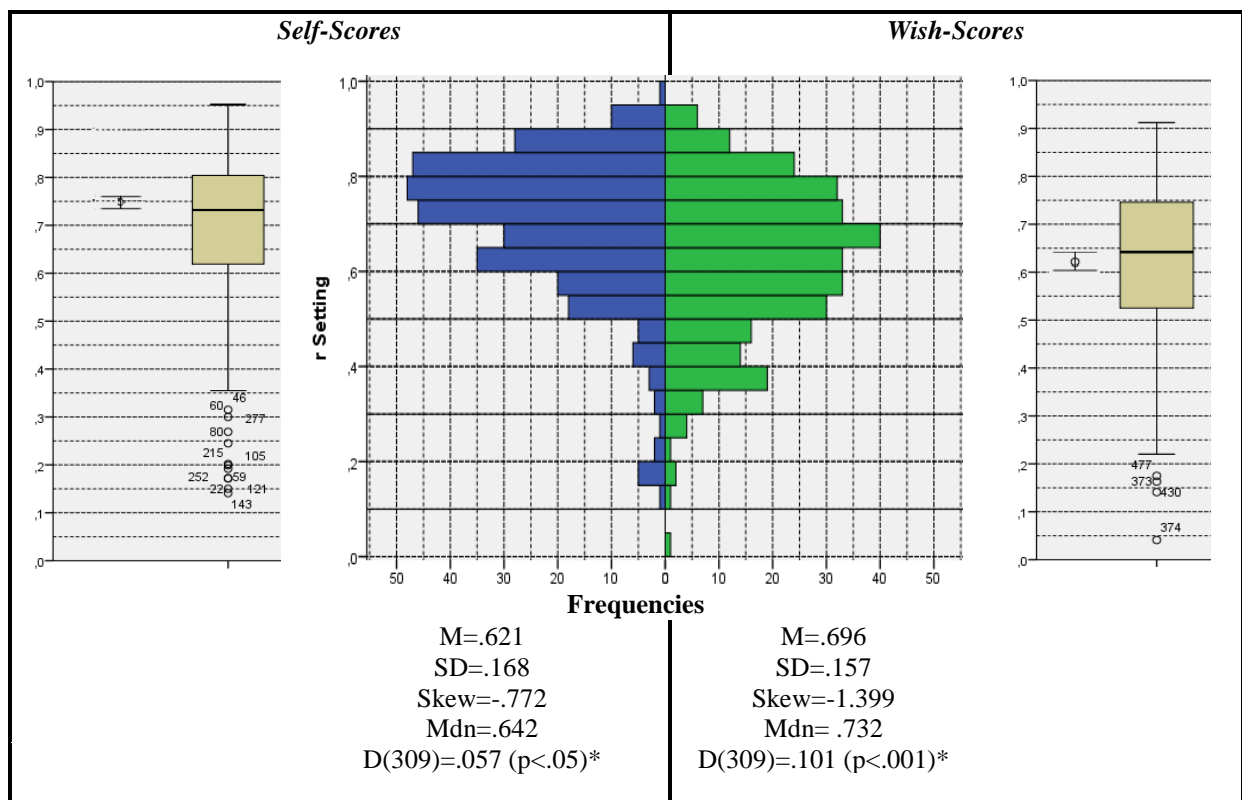
Figure 20. Distributions for individual setting-correlations (n=309).

Notes. For Wish-Scores and Self-Scores, answers between private and *public* settings were correlated per person. *Histograms* (*centre*), boxplots, averages and confidence interval of 95% (*margins*) are displayed. **p<.05 in Kolmogorov-Smirnov test for normal distribution.*

Significant Kolmogorov-Smirnov tests reflected that a normal distribution was unlikely for both setting-correlations and setting-differences. The MANOVA was used nonetheless, since it is the stricter method and very robust against infraction of distributional prerequisites. Individual setting-differences and setting-correlations all had negatively skewed distributions. Therefore, higher values were more likely than lower ones. Conspicuously, the ranges for distributions were much larger for *Wish-Scales* than for *Self-Scales*. Setting-differences and setting-correlations were not significantly correlated with one another (r(307)<|.08|, p>.18).

The inner subject MANOVA resulted in a significant main effect (F=(2,307)=37.799, p<.001): setting-differences were higher for *Self-Scores* than for *Wish-Scores* (F(1,308)=20.35, p<.001, ηp^2=.062), i.e., subjects *publicly* exaggerated their own characteristics more frequently than they did for their partner preferences. Moreover, *Self-Scores* were more congruent between the settings than *Wish-Scores* and, as a result, were affected by larger setting-distances *(*F(1,308)=49.87, p<.001, ηp^2=.139). Corresponding to this, the overall

higher setting-correlations for *Self-Scores* (r: m=.747, SD=.050) suggested that these were more robust against the setting change than the *Wish-Score*s (r: m=.795. SD=.042).

Due to the fundamental differences for answer biases between *Self-Scores* and *Wish-Scores*, all further calculations were run separately for these two instead of integrating the scores into one overall score.

6.3.2.2. Scores for faker detection

The less the answers of an individual varied in the *private* setting, the less they also varied in the *public* setting (*Wish-Score*s: r(307)=.825, p<.001; *Self-Score*s: r(307)=.811, p<.001). Furthermore, the higher the *private* scores were, the less the individual answers varied in *public* (*Wish-Score*s: r(307)=-.385, p<.001; *Self-Score*s: r(307)=-.394, p<.001). SDs of all *private* and *public* scores (*Self-Scales* and *Wish-Scales*) moderately correlated with the setting-correlations. *Table 28* shows the correlations to CVI and *public* SD[33].

Table 28: Pearson correlations of individual answer biases with faker scores (n=309)

Faker scores	M	SD	Setting-difference				Setting-correlation			
			Wish-Scores		Self-Scores		*Self-Scores*		*Wish-Scores*	
			r	p	r	p	r	p	r	p
Covariance Index CVI	.015	.074	.075	.188	-.015	.797	-.096	.091	.051	.373
Public SD *Self-Scales*	.986	.228	-.032	.574	-.052	.360	.237	<.001***	.385	<.001***
Public SD *Wish-Scales*	1.08	.253	.155	.006**	-.075	.190	.326	<.001***	.241	<.001***

Notes. *Individual answer biases in the form of differences and correlations between answers of public and private setting were separated for Wish-Scores and Self-Scores. **p<.01. ***p<.001*

The inner-subject ANOVA from the previous section also found a significant difference in the CVI between the settings (F(1, 308)=171.2, p<.001, ηp^2=.357, *Private*: M=-.0512, SD=.0894; *Public*: M=.0173, SD=.0736). Only seven subjects' public CVI scored higher than two SD's and were classified as 'fakers' (Chaney & Christiansen, 2004). *Public* CVI did not significantly correlate with any score for *answer bias*es (r(307)< |.10|, p>.09).

Moreover, no consistent pattern showed for *public* SD, though lower variations in the answers about own characteristics were slightly correlated with self-exaggeration (r(307)=.16, p<.01). Due to the lack of relevant higher correlations to the *answer biases,* the CVI and *public* SD's were not considered for further faker characterization or predictive validity subsequently.

---

[33] Appendix V.4 presents some histograms and further descriptive statistics.

6.3.2.3. Correlates to facets

No significant correlation between individual setting-differences to age, education, gender and relationship status occurred (r(307)<|.12| p>.05).

Table 29 presents the correlations between individual setting-differences to *Wish-Scales* and *Self-Scales*[34]. The following two examples explain the table:

- Individuals who generally scored higher *public* (m, r(307)=-.25, p<.001), and lower in *private Self-Scores* (m, r(307)=.27, p<.001), publicly more exaggerate answers about their own characteristics.

- Individuals who scored higher in privately preferred *physical attractiveness* (r(307)=.21, p<.001), or in *Socio-Economic Status* (r(307)=.17, p=.004), publicly less exaggerated their other preferences.

Table 29: Pearson correlations between setting differences to scales (n=309)

	Setting-differences							
	Wish-Scales				Self-Scales			
	Private		Public		Private		Public	
Scales	Self	Wish	Self	Wish	Self	Wish	Self	Wish
Score Average	.02	.26***	-.04	-.38***	.27***	-.04	-.25***	-.09
Standard Deviation	-.12	-.08	-.07	.16**	-.30***	-.05	-.05	-.03
Physical attractiveness	.00	.21***	-.02	-.02	.09	-.04	-.05	-.05
Sexuality	-.03	.12	-.01	-.12	.04	.02	-.16**	-.07
Gregariousness	-.01	.14	-.02	-.13	.14	-.15**	-.11	-.15
Socio-economic status	-.01	.17**	-.04	-.04	.24***	.06	.08	.01
Responsibility	.07	.12	-.04	-.11	.14	.01	-.04	-.09
Manners	-.02	.09	-.06	-.10	.10	-.11	-.11	-.11
Emotional stability	.01	.13	.00	-.13	.36***	.01	-.01	-.04
Extraversion	-.03	.11	-.05	-.13	.14	-.15**	-.06	-.16**
Openness	-.02	.12	-.07	-.13	-.05	.03	-.23*	-.03
Agreeableness	-.04	.14	-.05	-.21**	.06	-.06	-.19**	-.07
Conscientiousness	.05	.13	-.01	-.11	.17**	-.03	-.02	-.12

Notes. *The setting differences of answer scores between public and private answer setting were corrected by subtracting the corresponding scale value from the score before correlating.* ** p<.01.*** p<.001.

6.3.3. Effects on criterion-related validity

Table 30 shows that *private Self-Scales* were more strongly correlated with *relationship satisfaction* among couples than *public* ones, and that *private Wish-Scales* were less correlated with *relationship satisfaction* than *public* ones.[35]

---

[34] Appendix V.5 presents the correlations of both kind of answer biases (setting-differences and setting-distances) to (1) *Self-Scales* and *Wish-Scales* (2) demographics, (3), CVI (all n.s.) and answer SDs (all n.s.) with exact p values. In the present study only individual differences in answer shifts between settings were focused.

[35] Appendix V.6. presents correlations to *relationship satisfaction* with exact p values and an approach which applied a regression analysis and rejected CVI as a moderator for predicting RQ.

Table 30: Pearson correlations of scales with relationship satisfaction (n=225)

Scales	Setting	PAT	SEX	GAC	SES	RES	MAN	ES	E	O	A	C
Self	*Private*	.062	.062	.182**	.168*	.145*	.039	.263***	.181*	-.095	.022	.188**
	Public	.063	.011	.135*	.139*	.122	.004	.182**	.157*	-.047	-.015	.168*
Wish	*Private*	.087	-.023	.051	-.038	.115	.074	.042	.031	-.034	.091	.141*
	Public	.190**	.032	.044	.053	.138*	.124	.090	.050	-.037	.132*	.175**

*Notes. PAT: Physical Attractiveness. SEX: Sexuality. GAC: Gregariousness. SES: Socio-Economic Status. RES: Responsibility. MAN: Manners. N-: Emotional stability. E: Extraversion. O: Openness. A: Agreeableness. C: Conscientiousness.*p<.05. **p>.01. ***p<.001.*

6.4. Discussion

In this section, patterns of *answer biases* and approaches to handle these in mating contexts are discussed, linked to preceding research about *SDR,* and compared to those from job application contexts.

6.4.1. Main results

As a basis for evaluation and discussion, *Table 31* summarizes the core results of the present study corresponding to the hypotheses.

Table 31: Overview of key results of Study V

Content	Main results	Effect sizes
Hypothesis 1a: Exaggeration of socially desirable characteristics.	Most own characteristics were significantly exaggerated in the *public* self-description - in descending order: *emotional stability (+), physical attractiveness (-) gregariousness (+), agreeableness (+), extraversion (-), manners (+), conscientiousness (+), sexuality (+) and socio-economic status (-).*	Medium to small
	Only own *responsibility* (-) and *openness* (+) did not significantly shift between settings.	None
	Additionally, preferred s*ocio-economic status and physical attractiveness* were exaggerated (-) while the other assumed preferences for socially desirable partner characteristics were not (-).	Small None
Setting correlations	For most participants, setting-correlations of *Self-* and *Wish-Scores* were medium to high (r>.5). They were also high for all of the characteristics (0.85> r >0.63).	large
Hypothesis 1b: Differences for *Wish-Scales* and *Self-Scales*.	*Self-Scores* were more exaggerated than *Wish-Scores* (-).	Medium
	Self-Scales showed higher setting-correlations than *Wish-Scales* (-).	Large
	Most subjects (61%) exaggerated answers about their own personality, while only one third answered more modestly in public (20%). Around 20% did not shift their answers about preferences and also around 20% did not shift their answers about own characteristics (+). The same percentage of subjects significantly exaggerated and downgraded their partner preferences (both around 40%) (-).	
Hypothesis 1c: Gender effects	No relevant gender differences occurred for *answer biases* (+/-).	None

Content	Main results	Effect sizes
Hypothesis 2a: Faking scores	The CVI was not significantly associated with setting-differences or setting-correlations (-).	None
	The higher the SD of *public* answers was, the more differentiated the answer behavior and the higher was the setting-correlation for a person (+).	Medium
	No consistent correlation pattern of setting-differences to SD of *public* answers occurred (-).	None to small
Hypothesis 2b: Personality correlates	Self-exaggeration was associated with lower *private conscientiousness* (+) and *emotional stability* (-), higher *public agreeableness* (-) and *openness* (+) and higher preferred *Extraversion* (+), but not with own *Extraversion* (-). Additionally, it was linked to higher *public sexuality* and lower *private socio-economic status*. *Wish*-exaggeration was linked to lower *private*ly preferred *Socio-Economic Status* and *Physical Attractiveness* and higher *publicly* preferred *Agreeableness*.	Small, only ES medium
Hypothesis 3a: Predictive validity	Among couples, *private* answers about own characteristics predicted *relationship satisfaction* better than *public* ones. (+)	Small
	Among couples, *private* answers about partner preferences predicted *relationship satisfaction* worse than *private* ones. (+)	Small

Note. +*: evidence confirming the corresponding hypothesis. −: evidence against the corresponding hypothesis.*

6.4.2. Conclusion

6.4.2.1. Explaining social desirable responding

The moderate exaggeration for socially desirable own traits in the current online dating setting could be explained as follows. In situations of self-disclosure, a truth - attraction tension typically evolves. On the one hand, motivations to be authentic and accurate have an influence: one generally seeks to see oneself - and to be seen - as honest (Mazar & Ariely, 2006), to be understood and appreciated for oneself as one truly is, particularly in close relationships (Walther, Anderson, & Park, 1994; Swann, De la Ronde, & Hixon, 1994; Laurenceau, Barrett, & Pietromonaco, 1998).

On the other hand, simultaneously, motivations to give a good impression, to be liked by others, and to influence the outcomes of applicative situations to one's advantage affect answering behavior (Buller & Burgoon, 1996). The different aspects of this tension enhance or reduce the probabilities of *answer biases* in many *high-stake* settings (Gibbs, Ellison, & Heino, 2006). Existing research (Ellison, Heino, & Gibbs, 2006) detected that one strategy to resolve the pressures of this truth – attraction - tension in mating contexts is to present one's *ideal self,* more than one's *present self.* The *ideal self* contains true notions by subsuming one's potentials from the past, in the present and in the future (Higgins, 1987).

Corresponding research found that self-presentational decisions represent avoidance-avoidance situations for online daters: Completely telling the truth potentially threatens an attractive appearance, yet telling lies potentially damages one's self-view and relational

goals. Individuals then tend to make ambiguous statements that are neither true nor deceptive. Possibly, a range in answer scale points exists, within which participants choices apply well enough to feel honest for them (Bavelas, Black, Chovil, & Mullett, 1990). To conclude, one strategy of appealing to an attractive potential partner would be to present oneself as maximally positive in the range of what is still compatible with the own self-portrait.

6.4.2.2. Shifts between settings

Interestingly, the actor and partner characteristics which correlated the highest with RQ in *Study IV* - namely *physical attractiveness*, *agreeableness*, and *emotional stability* - were exaggerated most within the dating setting of the current study. This finding is congruent with previous research results on what is particularly socially desirable and attractive in mating contexts (Dyrenforth, Kashy, Donnellan, & Lucas, 2010; Luo & Zhang, 2009).

One possible explanation why *Self-Scales* were more exaggerated than *Wish-Scales* is that partner characteristics are socially desirable, yet the preferences for these characteristics are not. Preferred s*ocio-economic status* and *physical attractiveness* may have been exaggerated since the corresponding *private* scores were among the lowest and participants tried to adjust these in the form of setting a minimal requirement, i.e., s*ocio-economic status* and *physical attractiveness* seem to be less important than most other traits in a partner but still might work with a cut-off criterion. The corresponding selective strategy would be to deselect candidates who do not exceed a minimum of attractiveness and status in advance, while the selection among the remaining potential partners will be conducted upon other traits.

In the current study, the setting-differences for the *Big Five*-related constructs ($0.2 < d < 0.8$) were similarly large as the ones found in so-called fake-good studies applying a within-subject design in the context of job application ($0.4 < d < 0.7$) (O'Connell, Kung, & Tris, 2011). There were some congruencies in which traits are perceived as social desirable in job and partner application settings, but also some relevant differences. In job application contexts, as well as in the current dating context, *emotional stability* was most exaggerated, while *openness* had a similar subordinated ranking. As differences in job application contexts, *conscientiousness* was similarly exaggerated as *emotional stability*, while it seemed to be less important in the current dating context. This finding was inverted for *extraversion* and *agreeableness*.

6.4.2.3. Setting correlations

The substantial and high individual inter-setting correlations suggested a certain degree of robustness against setting changes and SDR. As suggested by former qualitative research (Ellison, Hancock, & Toma, 2012), the present quantitative findings indicate that individuals predominantly answer honestly and precisely about the different characteristics in online dating situations. Intentional *answer biases* including SDR are only one reason for the unexplained variance. Other sources might have been:

- Unsystematic errors: e.g., sudden distractions, inner struggles, other circumstances;
- Low commitment to answer correctly: e.g., due to the long survey or one's duty as a participating psychology student to complete the attendee hours as part of the curriculum;
- Systematic factors of the setting: e.g., filling out the same questionnaire twice.

Despite the high overall setting robustness, the large differences between *Self-Scores* and *Wish-Scores* were remarkable, i.e., preferences were more biased based on individually varying influences beyond mean-level changes. *Wish-Scores* may be less retest-reliable because the self-image is clearer than the image of a potential mate – possibly because the self-assessment is an evaluation of a real person while the assessment of an ideal partner more is an act of imagination.

6.4.2.4. Gender differences

The non-existence of gender differences for *answer biases* contradicted previous studies linked to the theory of evolution, which indicated that strategies to attract potential partners differ between men and women: e.g., assuming that women lure more with beauty-related and men more with status-related attributes (McGrew, 1995). The MANOVA showed some effects which did not cross the strict level of significance (p<.01), but would have with p<.05. This finding indicated that the gender differences in the current dataset were so small that they most likely lack practical relevance. The homogenous sample of young and educated participants could be the reason for this finding: these may have had more modern and more similar values which affected their answer behavior more than typical gender-related influences, which could cause larger differences in an older sample.

6.4.2.5. Personality correlates of answer biases

As with previous studies, the present examination replicated rather small correlations of *answer biases* to *personality traits* (Ones, Viswesvaran, & Reiss, 1996; Furnham, 1986;

Pauls & Stemmler, 2003). The finding that *conscientiousness* was associated lower *answer biases* went along with existing research on SDR in job contexts (Dilchert, Ones, Viswesvaran, & Deller, 2006): *Conscientiousness* may go along with a more conscientious self-related answer behavior in most contexts, including job and the partner application setting. Moreover, some other relations to motives of impression management were indicated:

(1) The higher an individual scores in own *sexuality, openness,* and *agreeableness (public),* the more they / she may feel the need to impress, attract, and exaggerate oneself in an online dating setting, possibly because these traits positively correlate with how important finding a partner is, which causes a larger motivation to attract;

(2) The lower an individual scores in *emotional stability* and *socio-economic status,* the more he/she might feel the need to give an honest or modest impression of the own person, possibly due to their lower ego and, thereby, a more restricted ideal self.

(3) Individuals exaggerate their answers about social desirable own traits differently when applying for a partner than when applying for a job - based their scores in *extraversion*, *emotional stability,* and *agreeableness*. Another possible interpretation is that the results are partly non-congruent with findings from other studies about personality correlates of SDR, as these studies mostly tested SDR with SDR-scales in between-subject designs, rather than with real individual setting differences. Therefore, something different might have been measured there which in return would have led to different relations to the same constructs.

Moreover, existing research found that *answer biases* partly depended on the job applied for, thus specific job-relevant attitudes are affected (Birkeland, Manson, Kisamore, Brannick, & Smith, 2006). In the dating context, *answer biases* might also depend on the kind of partner applied for. The present study supported this idea, in that answer *shifts* were significantly associated not only with own characteristics but also with partner preferences:

(1) Self-exaggeration was associated more with preferring extraverted partners: Individuals who more prefer extraverted partners - i.e., partners who more like to deal with others - may run a strategy to attract these by optimizing their own socially desirability, thinking that this is even more important for more social, extraverted subjects.

(2) Wish-exaggeration was associated less with preferring agreeable partners (*public*) and more with preferring *socio-economic status* and *physical attractiveness (private)* in a partner. A possible explanation is that a type of subjects performs a higher selective attraction strategy: these might less strongly prefer agreeable partners and instead more focus on the partner's *socio-economic status* and *physical attractiveness* and, as a result exaggerate these. Correspondingly, these both preferences were only significantly exaggerated ones in *public*.

6.4.2.6. *Faker scores*

Both faker scores (CVI and public SD) showed to be inappropriate for faker detection in the online dating setting.

The large differences between *high-stake* and non-*high-stake* setting, i.e. between the private and the public CVI, were unsurprising: these differences are the basis the CVI is built upon. However, the *answer biases* were not affected by the CVI, indicating its irrelevance for criterion-related validity. This finding is inconsistent with the former studies on the CVI (Christiansen, Robie, & Bly, 2005; Chaney & Christiansen, 2004). An explanation could be that the setting (application for a partner or a job) and, hence the motivations for SDR were very different.

The second kind of faking indexes - the *public* SDs - were moderately associated with *answer biases,* as they were in former studies (Ellingson, Sackett, & Hough, 1999; Christiansen, Goffin, Johnston, & Rothstein, 1994; Hough, 1998; Ones, Viswesvaran, & Reiss, 1996). Nonetheless, SD was highly correlated with: (1) the general height of answers and (2) SD in the *private* setting. Thus, public SD probably represents a characteristic of answer behavior itself, being a consequence of a restricted ending of the answer scale, rather than a proper indicator of faking behavior. The reason may be that individuals who already privately scored highest on many items *private*ly could not exaggerate these publicly anymore. Corresponding to this, the most probable reason why the scale for own *responsibility* was not exaggerated is that it already had the highest *private* parameters (RES, m=5.05, SD=.79).

6.4.2.7. *Criterion-related validity*

For job settings, meta-analyses indicated that faking had only a minimal impact on personality tests' criterion-related validity and performance outcomes (Dilchert, Ones, Viswesvaran, & Deller, 2006). By contrast, the current study found some possibly relevant differences when predicting RQ with data from the two different settings: *private Self-Scales* predicted *relationship satisfaction* slightly better than *public* ones. This finding indicates that *private* answers concerning own traits may be more accurate in general, and additionally undermine the predictive validity of currently applied personality tests in online dating settings.

One reason why *public* preferences predicted RQ better than *private* ones, could be that the individuals who lived in a partnership during the current study had higher preferences *publicly* when their partner was 'better' (e.g., more agreeable, more good-looking, more conscientious) and, thereby, already more exaggerated their 'real' *private* preferences. Thereby,

positive partner effects on the own RQ might have indirectly been measured. The finding of *Study II,* that partners who stay in their partnership decrease in their scale values for the trait *optimizing demanding* over time, supported this idea.

6.4.3. Limitations

In this section and *Table 32*, the limitations and benefits of the present study are juxtaposed and discussed.

Table 32: Evaluation of the Study V

	Benefits	Limitations
Comparability	+ The within-subject-design enabled setting comparisons for different characteristics and individuals. + Different measures of social desirable responding including individual *answer biases* were examined and compared. + The questionnaire enabled a comparison between own and preferred characteristics.	- The within-subject-design could have evoked repetition biases. - A critical, possibly biasing component in the instruction of *private* setting was identified: *"Most subjects tend to slightly overestimate their selves, which is totally normal. Consider this for your answers."*
Generalizability	+ The sample size was adequate and only affected by a small dropout. + A strict level of significance was used.	- The sample of young and highly educated Germans was very homogenous. - The data was collected in a faked, not in a real online dating setting.

6.4.3.1.   Sample

The adequate sample size, the low dropout and the strict level of significance strengthened the validity of the findings. Nonetheless, the homogeneous sample did not allow further generalization of the findings on other cultures or demographically different sections of population such as on older or less educated individuals. Further work should examine *answer biases* for other cohorts. Data from true online dating settings would be desirable, as with the current faked setting it remains unclear, how well the participants pretended to answer in an online dating setting.

6.4.3.2.   Study design

Commonly, the differences found in within-subject designs were larger than the ones found in between-subject designs when comparing *high-stake* and non-*high-stake* situations (Viswesvaran & Ones, 1999). Within-setting-designs are less affected by error variance due to individual differences, since the same sample is tested twice: this leads to higher comparability of the data. Furthermore, individual scores of *answer biases* between settings could be calculated and correlated with traits. A disadvantage is that individuals may answer different-

ly due to effects of fatigue and practice when taking the same test twice. Moreover, a critical component of *private* instruction was identified (see *Table 32*): the reminder of common self-exaggeration could either have had a demand or a suppressing effect on how much participants already (or not yet) exaggerate in the *private* setting.

6.4.4. Practical implications

When determining *answer biases* in mating contexts, the question arises, how one can deal with these in a real-life scenario in a way that predictive validity is not potentially impeded. The exaggerated *high-stake* scores might be handled by using, e.g., *high-stake* situation-specific appropriate norms (Ones & Viswesvaran, 2007)**.** Two different norms for scaling - one from the *high-stake* setting and one from an anonymous investigation - would be applied then, such as Beauducel & Kersting (2010) realized it for a job-related personality test. The results of the present study indicated that this method could be particularly suitable for own characteristics in online dating settings, as these were exaggerated statistically and practically significantly by most of the subjects. For partner matching tests in an online dating setting (e.g., for the YOUME-L as a potential candidate), data from online daters would have to be collected to develop such appropriate norms. Another approach could be to correct scales by the average of the expected, common shifts between settings when a test is applied in an online dating context. A disadvantage of this approach is that it does not consider the deviating SDs of answers for the different settings.

A promising approach to analyze the sources of *answer biases* could be to qualitatively examine reasons for individual *shifts* first, e.g., with interviews using the method 'thinking loudly', while answering in an online dating setting. This might be a step to developing a more proper approach to understand, predict, and finally correct individual SDR. Moreover, future work has to show if, and how much, *answer biases* actually influence the predictive validity of models predicting RQ in a matching scenario. These elaborations could relevantly profit from methods of ML.

7.   *General Conclusion and Contribution*

In summary, the present work fundamentally has contributed to the understanding of how general and relationship-related personality traits interact with romantic relationships, mainly with their quality. The main insights refer to three different fields (content, method, applicability) and are outlined in this chapter.

7.1.   *Content: who makes whom happy in a romantic relationship?*

The results of the four-year longitudinal *Study I* answered the core question of the present work, supplemented by the results of the cross-sectional *Study IV* – with the help of three different personality tests for self-assessment (BB-PI, PD-I, YOUME-L). When examining linear additive models, it was predominantly the actors who made themselves happy in a relationship (*Study I* and *IV*). Over time, aspects – particularly those referring to more *relationship-related personality* as general competency in partnerships, conflict-related traits, *agreeableness*, and *emotional stability* - decided upon *relationship satisfaction*, separation (intents), *conflicts,* and *harmony* (*Study I*). The present work replicated previous research in the field by finding that similar partners were happier with one another: This was shown by direct correlations (*Study I* and *III*) and, for the first time, by the weights of similarity variables in cross-validated prediction models, which exceeded 50% of the overall variable importance (*Study I* and *IV*).

Moreover, it could be shown for the first time, how varying constellations of both partner's values affect RQ differently for diverse traits (*Study I* and *III*). To that, a novel approach to score (dis)similarities was applied. Results indicated two patterns: (1) Similarities in socially desirable traits more often affected RQ positively when both partners scored high, and more often negatively, when both scored low – this was inverse for socially undesirable traits; (2) For ideal-real fittings, in many cases, partner preferences seemed to activate the effect of the preferred trait - if the actor highly preferred a trait, high partner values in this trait were often beneficial and low ones restraining, but when the actor did not highly prefer this trait in a partner, RQ was merely affected (*Study III*). Accordingly, for preference-related traits, such as 'need for closeness' or 'faithfulness', the constellation of a high actor score and a low partner score seemed to especially affect RQ negatively (*Study I*).

With some remaining relevant, unsolved issues and limitations, the BB-PI (*Study I* and *II*) and the more economic YOUME-L (*Study III, IV* and *V*) were shown to be promising

candidates for application in relevant mating contexts, e.g., as matching tests for online dating websites.

7.2.   Methods: how can the prediction models be improved?

The present work presents unprecedentedly reproducible results of personality models, which significantly predicted RQ in a CV (*Study I* and *IV*): with the help of methods from ML, 37% of the variance could be explained after four years. This largely exceeds the moderate correlations that previous research without ML has replicated thus far (*Study I*).

ML was tested and validated as an appropriate method to predict social outcomes, based on so-called 'soft' data, i.e., self-assessed own personality ratings. To conclude, these methods can contribute to the economy of psychological tests, which are supposed to be applied in real-life settings: On the one hand, the highly efficient methods of integrated variable selection can contribute to reduce the lengths of tests. On the other hand, all the mechanisms from ML, which intent to reduce the overfitting of models, allow a more accurate determination and evaluation of the additive predictive test validity (*Study I* and *IV*). The additive predictive validity is useful beyond the predictive validities for single scales, especially regarding tests, which are designed for a specific predictive purpose.

A novel matching test - the so-called the YOUME-List - was designed and applied in the present work: it integrates the same scale contents for preferred partner characteristics, and for own characteristics. The predictive scale validities for preferences and fittings were as high as those for similarities (*Study III*), but they did not additionally contribute to the cross-sectional model optimization (*Study IV*).

The relevance of contents was analyzed with different predictive models. In the longitudinal predictions with the BB-PI and PDI (*Study I*), actor effects turned out to explain nearly all possible variance of RQ on their own (33 of 37% max.), while partner, similarity and hypotheses-based moderator effects did not significantly incrementally contribute. In the cross-sectional predictions with the BB-PI and PDI, none of these variable domains could relevantly explain variance of the RQ (<1%). Contrarily, in the cross-sectional predictions with the YOUME-L (*Study IV*), the actor-partner-interaction effects – e.g., similarities - explained the most of the variance of RQ: the best model (19%) was based on these hypotheses-based moderators only. This gave incentives for future work, to examine interaction effects based on the YOUME-L more closely, especially as actor-partner-interactions could potentially contribute to match different personalities.

S*tudy I* and *IV* followed and supported the current, but unfortunately slight, trend in social and cognitive psychology to broaden their data-analytical methods. The findings supported that the best approach to explain a phenomenon might be to accurately predict it. Thereby, psychology as a natural science still in its infancy, has made a further step in the more methodically and statistically sophisticated direction of the 'grown-up' sciences, such as physics and computer science. The state-of-the-art models within these fields more accurately predict outcomes through interacting variables while still managing to explain the function behind these interactions (Chung, Senior, Vinyals, & Zisserman, 2017; LeCun, Bengio, & Hinton, 2015; Vinyals, Toshev, Bengio, & Erhan, 2015). A feasible strategy to induce a methodological revolution and to better establish the psychology as a natural science could lie in a more interdisciplinary transference of knowledge and know-how from these fields.

7.3. *Applicability: what restricts generalizability to real-life mating contexts?*

The transferability of models is an essential prerequisite for the application of *relation-related personality* tests in mating contexts. In the present studies, which focused directly on predictive accuracy, two major limitations in generalizability were present:

(1) *Study I* and *IV* showed that cross-sectional predictions perform differently than longitudinal forecasts. Not only were they less precise, but also the personality domains that turned out to be relevant differed immensely. Therefore, partner matching models in real mating contexts should always be created using longitudinal sections to ensure long-term predictive validity.

(2) On the same cross-sectional dataset from *Study III* and *IV*, large cultural differences between Persian and German couples in predictive validity were found: Models for partner matching in real mating contexts should be specifically developed for the culture they are meant to be applied in, as they cannot be transferred safely.

Two additional studies directly focused on possible constraints, which may affect transferability and predictive accuracy when the models of *Study I* or *IV* would be applied in a dating context:

(1) In *study V, answer biases* as SDR were examined in a within-subject design including an anonymous and a faked online dating setting. High inter-setting correlations for subjects, as well as for all traits, supported a general context-overarching measurement accuracy of self-assessed own and preferred *personality traits*. An overall tendency for self-

exaggerating and individually differing answer distortion tendencies were found, yet an attempt for individual faker-detection based on answer variances failed. The anonymous and online-dating answers slightly differed in their predictive validity for RQ: this emphasized a possible practical relevance of *answer biases* in mating contexts.

(2) In *Study II*, the stability of *relationship-related personality traits* over time and over different relationship statuses was examined to evaluate the patterns of relationship-personality interaction. Aside from the generally moderate retest-retest reliabilities for facets over the four years, current partnerships correlated with the changeability of *relationship-related personality*. This novel finding expanded the few existing inconsistent previous research results (Caspi & Herbener, 1990; Kirkpatrick & Hazan, 1994; Scharfe & Bartholomew, 1994). Particularly, RQ-related changeability potentially disturbs the transferability of models based on partners who already influenced one another to singles who not yet know one another. Due to their association pattern with the current relationship status – and possibly confounding with RQ -, the suitability for mating contexts is particularly questionable for facets linked to *dominance* and *submission* and *tendencies to react emotionally positive or negative.* However, all sex-related traits were robust against the current relationship experiences and, hence, were evaluated as highly applicable for matching models from the perspective of reliability and robustness.

In Study II and *V*, an approach was proposed to handle mean-level differences, when *personality traits* are self-assessed in a mating setting: the scale values can be corrected by the expected average of the typical trait-, person- and surrounding-specific decrease or increase.

The present work tested a broad range of psychometric requirements, which exceeded the level of a common test evaluation by analyzing cross-contextual applicability (cross-relationship, cross-setting, and across time). This 'additional' quality criterion matters greatly, should the corresponding diagnostic instrument be applied in one specific, or even in different, contexts that exist beyond research studies. Some of the insights gained lay the foundation on which questionnaires for self-assessment of personality traits can be responsibly applied and interpreted in the real mating context. Thereby, the results of the present work showed several reasons why most known and current matching tests do not yet seem to bring together couples who are happier with each other than with the free selection. (Cacioppo, Cacioppo, Gonzaga, Ogburn, & VanderWeele, 2013).

Literature

Andresen, B. (2012). *Beziehungs- und Bindungs-Persönlichkeitsinventar. Manual. (BB-PI) [Relationship- and Attachment-related Personality Inventory. Manual].* Göttingen: Hogrefe.

Andresen, B. (2012). Wunschpartner-Realpartner-Listen (WPRP-L) [Wished Partner- Real Partner- Lists]. In B. Andresen, *Beziehungs- und Bindungs-Persönlichkeitsinventar. Manual. (BB-PI) [Relationship- and Attachment-related Personality Inventory. Manual]* (pp. 104-105, Appendix). Göttingen: Hogrefe.

Andresen, B. (2015). Dreizehn-Persönlichkeitsdomänen-Inventar [Thirteen Personality Domain Inventory]. In B. Andresen, *Mythos Big Five - Neue Basisdimensionen der Persönlichkeit* (pp. 331-401, Chapter 5). Norderstedt: Books on Demand.

Barelds, D. P. (2005). Self and partner personality in intimate relationships. *European Journal of Personality, 19(6)*, pp. 501–518.

Bargh, J. A., McKenna, K. Y., & Fitzsimons, G. M. (2002). Can you see the real me? Activation and expression of the "true self" on the Internet. *Journal of social issues, 58(1)*, pp. 33-48.

Barrick, M. R., Mount, M. K., & Judge, T. (2001). Personality and performance at the beginning of the new millennium:What do we know and where do we go next? *International Journal of Selection and Assessment, 9(1-2)*, pp. 9-30.

Bartholomew, K., & Horowitz, L. (1991). Attachment styles among young adults: a test of a four-category model. *Journal of Personality and Social Psychology, 61(2)*, pp. 226-244.

Bavelas, J. B., Black, A., Chovil, N., & Mullett, J. (1990). Equivocal communication. In *Sage series in interpersonal communication, 11*. Thousand Oaks, CA, US: Sage Publications.

Beauducel, A., & Kersting, M. (2010). *Testbatterie für Berufseinsteiger-Persönlichkeit [Battery to test personality of career starters] (START-P).* Hogrefe.

Becker, O. A. (2012). Effects of similarity of life goals, values, and personality on relationship satisfaction and stability: Findings from a two-wave panel study. *Personal Relationships, 20(3)*, pp. 443–461.

Berry, C. M., Ones, D. S., & Sackett, P. R. (2007). Interpersonal deviance, organizational deviance, and their common correlates. *A review and meta-analysis. Journal of Applied Psychology, 92(2)*, pp. 410-424.

Berscheid, E. (1999). The greening of relationship science. *American Psychologist 54(4)*, pp. 260-66.

Birkeland, S. A., Manson, T. M., Kisamore, J. L., Brannick, M. T., & Smith, M. A. (2006). A meta-analytic investigation of job applicant faking on personality measures. *International Journal of Selection and Assessment*, pp. 317-335.

Botwin, M. D., Buss, D. M., & Shackelford, T. K. (2006). Personality and mate preferences: Five factors in mate selection and marital satisfaction. *Journal of Personality, 65(1)*, pp. 107-136.

Bouckaert, R. R., & Frank, E. (2004). Evaluating the replicability of significance tests for comparing learning algorithms. *In Pacific-Asia Conference on Knowledge Discovery and Data Mining*, pp. 3-12.

Brähler, E., & Brähler, C. (1993). *Paardiagnostik mit dem Gießen-Test [Couple diagnostics with Gießen-Test]*. Göttingen: Hogrefe.

Brym, R. J., & Lenton, R. L. (2001, February 6). *Love online: A report on digital dating in Canada.* Toronto, US: MSN.ca.

Buller, D. B., & Burgoon, J. K. (1996). Interpersonal deception theory. *Communication theory, 6(3)*, pp. 203-242.

Buss, D. M. (1989). Sex differences in human mate preferences: Evolutionary hypotheses tested in 37 cultures. *Behavioral and brain sciences, 12(1)*, pp. 1-14.

Cacioppo, J. T., Cacioppo, S., Gonzaga, G. C., Ogburn, E. L., & VanderWeele, T. J. (2013). Marital satisfaction and break-ups differ across online and offline meeting venues. *Proceeding of the National Acedemy of Science of the United States of Amarica, 110(25)*, pp. 10135–10140.

Cairns, R. B., & Hood, K. E. (1983). Continuity in social development: A comparative perspective on individual difference prediction. In P. B. Baltes, & O. G. Brim, *Life-span development and behavior* (pp. 301-358). New York, US: Academic Press.

Caspi, A., & Herbener, E. S. (1990). Continuity and change: assortative marriage and the consistency of personality in adulthood. *Journal of personality and social psychology, 58(2)*, pp. 250-258.

Caspi, A., & Roberts, B. W. (1999). Personality continuity and change across the life course. In L. A. Pervin, & O. P. John, *Handbook of personality psychology: Theory and research* (pp. 300-326). New York, US: Guilford Press.

Cawley, G. C., & Talbot, N. L. (2010). On over-fitting in model selection and subsequent selection bias in performance evaluation. *Journal of Machine Learning Research*, pp. 2079-2107.

Chaney, S., & Christiansen, N. D. (2004). Disentangling socially desirable responding from personality: Using covariance to detect response distortion. *In 19th Annual Meeting of the Society for Industrial and Organizational Psychology.* Chicago, US: IL.

Christiansen, N. D., Goffin, R. D., Johnston, N. G., & Rothstein, M. G. (1994). Correcting the 16PF for faking: effects on criterion-related validity and individual hiring decisions. *Personnel Psychology, 47(4)*, pp. 847-860.

Christiansen, N. D., Robie, C., & Bly, P. R. (2005). Using covariance to detect applicant response distortion of personality measures. *Symposium conducted at the 20th Annual Conference for the Society of Industrial and Organizational Psychology. Faking research: New methods, new samples, and new questions.* Los Angeles, CA, US.

Chung, J. S., Senior, A., Vinyals, O., & Zisserman, A. (2017). Lip reading sentences in the wild. *Computer Science. Computer Vision and Pattern Recognition. arXiv:1611.05358 [cs.CV]*, pp. 1-12.

Costa Jr, P., Terracciano, A., & McCrae, R. R. (2001). Gender differences in personality traits across cultures: robust and surprising findings. *Journal of Personality and Social Psychology, 81(2)*, pp. 322-331.

Couch, D., & Liamputtong, P. (2008). Online dating and mating: The use of the internet to meet sexual partners. *Qualitative Health Research, 18(2)*, pp. 268-279.

Dietterich, G. (1998). Approximate statistical tests for comparing supervised classification learning algorithms. *Neural Computation, 10(7)*, pp. 1895–1924.

Dilchert, S., Ones, D. S., Viswesvaran, C., & Deller, J. (2006). Response distortion in personality measurement: born to deceive, yet capable of providing valid self-assessments. *Psychology Science, 3*, pp. 209-225.

Domingos, P. (2012). A few useful things to know about machine learning. *Communications of the ACM, 55(10)*, pp. 78–87.

Donnellan, M., Larsen-Rife, D., & Conger, R. (2005). Personality, family history, and competence in early adult romantic relationships. *Journal Personality & Social Psychology, 88*, pp. 562-76.

DrexelUniversity. (2012, January 12). *ScienceDaily. Your source for the latest research news*. Retrieved September 5, 2014, from Looking for love: Researchers put online dating to the test: http://www.sciencedaily.com/releases/2012/01/120105142811.htm

Dyrenforth, P. S., Kashy, D. A., Donnellan, M. B., & Lucas, R. E. (2010). Predicting relationship and life satisfaction from personality in nationally representative samples from three countries: The relative importance of actor, partner, and similarity effects. *Journal of Personality and Social Psychology, 99(4)*, pp. 690-702.

Eastwick, P. W., Luchies, L. B., Finkel, E. J., & Hunt, L. L. (2014). The predictive validity of ideal partner preferences: A review and meta-analysis. *Psychological Bulletin, 140(3)*, pp. 623-665.

Ellingson, J. E., Sackett, P. R., & Hough, L. M. (1999). Social desirability corrections in personality measurement: Issues of applicant comparison and construct validity. *Journal of Applied Psychology, 84(2)*, pp. 155-166.

Ellison, N. B., Hancock, J. T., & Toma, C. L. (2012). Profile as promise: A framework for conceptualizing veracity in online dating self-presentations. *New Media Society, 14(1)*, pp. 45–62.

Ellison, N., Heino, R., & Gibbs, J. (2006). Managing impressions online: Self-presentation processes in the online dating environment. *Journal of Computer-Mediated Communication, 11(2)*, pp. 415-441.

Erol, R. Y. (2014). Development of self-esteem and relationship satisfaction in couples: Two longitudinal studies. *Developmental Psychology, 50(9)*, pp. 2291-2303.

Fahrenberg, J. M., M., S. J., & Brähler, E. (2000). *Fragebogen zur Lebenszufriedenheit [Life Satisfaction Questionnaire] (FLZ).* Göttingen: Hogrefe.

Feingold, A. (1992). Gender differences in mate selection preferences: A test of the parental investment model. *Psychological Bulletin, 112(1)*, pp. 125-139.

Findeisen, K. (2014). *Psychometrische Validierung eines Interaktionsfragebogens auf der theoretischen Grundlage der Gottman'schen Paartypen [Psychometric validation of an interaction questionnaire on the theoretical basis of Gottman's types of couples] (Bachelor thesis).* Hamburg, Germany: University of Hamburg.

Finkel, E. J., Eastwick, P. W., Karney, B. R., Reis, H. T., & Sprecher, S. (2012). Online dating: A critical analysis from the perspective of psychological science. *Psychological Science in the Public Interest, 13(1)*, pp. 3-66.

Fletcher, G., Simpson, J., & Thomas, G. (2000). Ideals, perceptions and evalations in early relationship development. *Journal of Personality and Social Psychology, 79(6)*, pp. 933-940.

Fowers, B. J., & Olsen, D. H. (2007). Predicting marital success with prepare: a predictive validity study. *Journal of Marital and Family Therapy, 12(4)*, pp. 403-413.

Furnham, A. (1986). Response bias, social desirability and dissimulation. *Personality and individual differences, 7(3)*, pp. 385-400.

Gaunt, R. (2006). Couple similarity and marital satisfaction: Are similar spouses happier? *Journal of Personality 74(5)*, pp. 1401-1420.

Gelman, A., & Loken, E. (2013). The garden of forking paths: Why multiple comparisons can be a problem, even when there is no "fishing expedition" or "p-hacking" and the research hypothesis was posited ahead of time. *Department of Statistics, Columbia University*.

George, D., Luo, S., Webb, J., Pugh, J., Martinez, A., & Foulston, J. (2015). Couple similarity on stimulus characteristics and marital satisfaction. *Personality and Individual Differences, 86*, pp. 126–131.

Gerdvilyte, A., & Abhyankar, S. C. (2010). The compatibility of ideal and real romantic partner characteristics, attachment to partner and relationship satisfaction among indian women. *Psychological Studies, 55(3)*, pp. 188-194.

Gibbs, J. L., Ellison, N. B., & Heino, R. D. (2006). Self-presentation in online personals the role of anticipated future interaction, self-disclosure, and perceived success in Internet dating. *Communication Research, 33(2)*, pp. 152-177.

Gottman, J. M. (2014). *What predicts divorce?: The relationship between marital processes and marital outcomes.* New York, US: Psychology Press.

Gottman, J. M., & Notarius, C. I. (2000). Decade review: Observing marital interaction. *Journal of Marriage and the Family, 62(4)*, pp. 927–947.

Hahlweg, K. (1996). *Fragebogen zur Partnerschaftsdiagnostik (…). Questionnaire for partnership diagnostics (FPD). [Partnership Questionnaire PFB, Problem List PL and Questionnaire on Life History and Partnership FLP].* Göttingen: Hogrefe.

Harvey, J. H., Wenzel, A., & Sprecher, S. (2006). The handbook of sexuality in close relationships. *Archives of Sexual Behavior, 35(6)*, pp. 759–760.

Hazan, C., & Shaver, P. (1987). Romantic love conceptualized as an attachment process. *Journal of Personality and Social Psychology, 52(3)*, pp. 511-524.

Heino, R. D., Ellison, N. B., & Gibbs, J. L. (2010). Relationshopping: Investigating the market metaphor in online dating. *Journal of Social and Personal Relationships, 27(4)*, pp. 427-447.

Henrich, G., & Herschbach, P. (2000). Questions on Life Satisfaction (FLZ M): A short questionnaire for assessing subjective quality of life. *European Journal of Psychological Assessment, 16(3)*, pp. 150-159.

Herbenick, D., Reece, M., Schick, V., Sanders, S. A., Dodge, B., & Fortenberry, J. D. (2010). Sexual behavior in the United States: Results from a national probability sample of men and women ages 14–94. *Journal of Sexual Medicine, 7(5)*, pp. 255–265.

Hicks, M. W., & Platt, M. (1970). Marital happiness and stability: A review of the research in the sixties. *Journal of Marriage and Family, 32(4)*, pp. 553-574.

Higgins, E. T. (1987). Self-discrepancy: a theory relating self and affect. *Psychological review, 94(3)*, pp. 319-340.

Hitsch, G. J., Hortaçsu, A., & Ariely, D. (2010). Matching and sorting in online dating. *The American Economic Review, 100(1)*, pp. 130-163.

Hofstede, G. (2001). *Culture's Consequences: comparing values, behaviors, institutions, and organizations across nations (2nd ed.).* Thousand Oaks, CA, US: SAGE Publications.

Hough, L. M. (1998). Effects of intentional distortion in personality measurement and evaluation of suggested palliatives. *Human Performance, 11(2-3)*, pp. 209-244.

Hudson, N., & Fraley, R. C. (2014). Partner similarity matters for the insecure: attachment orientations moderate the association between similarity in partners' personality traits and relationship satisfaction. *Journal of Research in Personality, 53*, pp. 112–123.

Humor Rainbow, I. (2015). *okcupid.* Retrieved February 2015, from Calculating Match Percentages: https://www.okcupid.com/help/match-percentages

Joel, S., Eastwick, P., & Finkel, E. J. (2017). *Is Romantic Desire Predictable? Machine Learning Applied to Initial Romantic Attraction. Retrieved from osf.io/gu8z7.* Retrieved from osf.io/gu8z7

Kaiser, A. L. (2014). *Partnerpassung in Beziehung: Einfluss gewünschter Partnermerkmale und eigener Merkmale auf die Beziehungsqualität und partnerschaftliche Probleme [The influence of preferred partner and own traits on relationship quality and conflicts] (Master thesis).* Hamburg, Germany: University of Hamburg.

Karney, B. R., & Bradbury, T. N. (1995). The longitudinal course of marital quality and stability: A review of theory, methods, and research. *Psychological Bulletin, 118(1)*, pp. 3-34.

Keizer, R., & Komter, A. (2015). Are "equals" happier than "less equals"? A couple analysis of similarity and well-being. *Journal of Marriage and Family, 77(4)*, pp. 954–967.

Kiecolt-Glaser, J. K., & Newton, T. L. (2001). Marriage and health: His and hers. *Psychological Bulletin, 127(4)*, pp. 472-503.

Kirkpatrick, L. A., & Hazan, C. (1994). Attachment styles and close relationships: A four-year prospective study. *Personal Relationships, 1(2)*, pp. 123-142.

Kohn, J. L., Rholes, S. W., Simpson, J. A., Martin A., M., Tran, S., & Wilson, C. L. (2012). Changes in marital satisfaction across the transition to parenthood. The role of adult attachment orientations. *Personality and Social Psychology Bulletin, 38(11)*, pp. 1506-1522.

Kuhn, M. (2016). *The caret Package*. Retrieved April 14, 2017, from https://topepo.github.io/caret

Laurenceau, J.-P., Barrett, L. F., & Pietromonaco, P. R. (1998). Intimacy as an interpersonal process: the importance of self-disclosure, partner disclosure, and perceived partner responsiveness in interpersonal exchanges. *Journal of personality and social psychology, 74(5)*, pp. 1238-1251.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521 (7553)*, pp. 436-444.

Lee, J. (1988). Love-Styles. In R. Sternberg, & M. L. Barnes, *The psychology of love* (pp. 38-67). New Haven, CT, US: Yale University Press.

Lev-Wiese, R., & Al-Krenawi, A. (1999). Attitude towards marriage and marital quality: A comparison among Israeli Arabs differentiated by religion. *Family Relations, 48(1)*, pp. 51-56.

Lucas, R. E., & Donnellan, M. B. (2013). Improving the replicability and reproducibility of research published in the Journal of Research in Personality. *Journal of Research in Personality, 4(47)*, pp. 453-454.

Luo, S. (2009). Partner selection and relationship satisfaction in early dating couples: The role of couple similarity. *Personality and Individual Differences, 47(2)*, pp. 133–138.

Luo, S., & Zhang, G. (2009). What leads to romantic attraction: Similarity, reciprocity, security, or beauty? Evidence from a speed-dating study. *Journal of Personality, 77(4)*, pp. 933–964.

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: A meta-analysis. *Journal of Research in Personality, 44(1)*, pp. 124-127.

Mazar, N., & Ariely, D. (2006). Dishonesty in everyday life and its policy implications. *Journal of Public Policy & Marketing, 25(1)*, pp. 117-126.

McGrew, W. (1995). Sex and age differences in preferences and tactics of mate attraction: Analysis of published advertisements. *Ethology and Sociobiology, 15(2)*, pp. 59-72.

Meltzer, A. L., McNulty, J. K., Jackson, G. L., & Karney, B. R. (2014). Sex differences in the implications of partner physical attractiveness for the trajectory of marital satisfaction. *Journal of Personality and Social Psychology, 106(3)*, pp. 418-428.

Minkov, M. (2007). *What makes us different and similar: A new interpretation of the World Values Survey and other cross-cultural data.* Sofia, Bulgaria: Klasika y Stil Publishing House.

Mischel, W., Ayduk, O., Berman, M. G., Casey, B. J., Gotlib, I. H., Jonides, J., . . . Shoda, Y. (2010). 'Willpower' over the life span: decomposing self-regulation. *Social Cognitive and Affective Neuroscience, 6(2)*, pp. 252–256.

Montoya, R. M., Horton, R. S., & Kirchner, J. (2008). Is actual similarity necessary for attraction? A meta-analysis of actual and perceived similarity. *Journal of Social and Personal Relationships, 35 (6)*, S. 889-922.

Moss, H. A., & Susman, E. J. (1980). Longitudinal study of personality development. In O. G. Brim, & J. Kagan, *Constancy and change in human development* (pp. 530-595). Cambridge, MA, US: Harvard University Press.

Nadeau, C., & Bengio, Y. (2003). Inference for the generalization error. *Machine Learning, 52*, pp. 239-281.

Noftle, E. E., & Shaver, P. R. (2006). Attachment dimensions and the big five personality traits: Associations and comparative ability to predict relationship quality. *Journal of Research in Personality, 40(2)*, pp. 179–208.

O'Connell, M. S., Kung, M.-C., & Tris, E. (2011). Beyond impression management: Evaluating three measures of response distortion and their relationship to job performance. *International Journal of Selection and Assessment, 19(4)*, pp. 340-351.

Ones, D. S., & Viswesvaran, C. (2007). Labor market influences on personality scale scores among job applicants. Four field studies in personnel selection settings. *Zeitschrift für Personalpsychologie [Journal for Personnel Psychology], 6(2)*, pp. 71-84.

Ones, D. S., Viswesvaran, C., & Reiss, A. D. (1996). Role of social desirability in personality testing for personnel selection: the red herring. *Journal of Applied Psychology, 81(6)*, pp. 660-679.

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science, 349(6251)*, pp. 1-8.

Orth, U. (2013). How large are actor and partner effects of personality on relationship satisfaction? The importance of controlling for shared method variance. *Personality and Social Psychology Bulletin, 39(10)*, pp. 1359-1372.

Ostendorf, F., & Angleitner, A. (2003). *NEO-Persönlichkeitsinventar nach Costa und McCrae, Revidierte Fassung. Manual [NEO-Personality Inventory according to Costa and McCrae, Revised Version] (NEO-PI-R).* Göttingen: Hogrefe Verlag.

Paki, S. (2014). *Wie gut passen iranische und deutsche Paare anhand ihrer persönlichen Wünsche und Merkmale zueinander? [How well do Iranian and German couples match one another based on their preferences and characteristics?] (Diploma thesis).* Hamburg, Germany: University of Hamburg.

Pauls, C. A., & Stemmler, G. (2003). Substance and bias in social desirability responding. *Personality and Individual Differences, 35(2)*, pp. 263-275.

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)*, pp. 576-593.

Regan, P. C., Levin, L., Sprecher, S., Christopher, S., & Gate, R. (2000). Partner preferences. What characteristics do men and women desire in their short-term sexual and long-term romantic partners? *Journal of Psychology & Human Sexuality, 12(3)*, pp. 1-21.

Robert Böhm, A. S. (2010). Are we looking for positivity or similarity in a partner's outlook on life? Similarity predicts perceptions of social attractiveness and relationship quality. *The Journal of Positive Psychology, 5(6)*, pp. 431-438.

Roberts, B. W., & DelVecchio, W. F. (2000). The rank-order consistency of personality traits from childhood to old age: A quantitative review of longitudinal studies. *Psychological Bulletin, 126(1)*, pp. 3-25.

Roberts, B. W., Helson, R., & Klohnken, E. C. (2002). Personality development and growth in women across 30 years: Three perspectives. *Journal of Personality, 70(1)*, pp. 79-102.

Roberts, B. W., Walton, K. E., & Viechtbauer, W. (2006). Patterns of mean-level change in personality traits across the life course: A meta-analysis of longitudinal studies. *Psychological Bulletin, 132(1)*, pp. 1-25.

Robins, R. W., Caspi, A., & Moffitt, T. E. (2002). It's not just who you're with, it's who you are: personality and relationship experiences across multiple relationships. *Journal Personal, 70(6)*, pp. 925-1064.

Rosenfeld, A., Zuckerman, I., Azaria, A., & Kraus, S. (2012). Combining psychological models with machine learning to better predict people's decisions. *Synthese, 189(1)*, pp. 81-93.

Rost, J. (2004). *Lehrbuch Testtheorie, Testkonstruktion [Textbook Test Theory, Test Construction]* (2 ed.). Bern: Huber.

Rowatt, W. C., Cunningham, M. R., & Druen, P. B. (1999). Lying to get a date: the effect of facial phsical attractiveness on the willingness to deceive prospective dating partners. *Journal of Social and Personal Relationships, 16(2)*, pp. 209-223.

Rudman, L. A., & Phelan, J. E. (2007). The interpersonal power of feminism: Is feminism good for romantic relationships? *Sex Roles, 57(11-12)*, pp. 787–799.

Ruvolo, A. P., & Veroff, J. (1997). For better or for worse: Real-ideal discrepancies and the marital well-being of newlyweds. *Journal of Personal and Social Relationships, 14(2)*, pp. 223-242.

Scharfe, E., & Bartholomew, K. (1994). Reliability and stability of adult attachment patterns. *Personal Relationships, 1(1)*, pp. 23-43.

Schmitt, D., & Buss, D. (2000). Sexual dimensions of person description: beyond or subsumed by the Big Five. *Journal of Research in Personality, 34(2)*, pp. 141-177.

Shaver, P. R. (1992). Attachment styles and the "Big Five" personality traits: Their connections with each other and with romantic relationship outcomes. *Social Psychology Bulletin, 18(5)*, pp. 536-545.

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science, 22(11)*, pp. 1359-1366.

Smolic, E. (2012). *Beziehungspersönlichkeit und -qualität (...) [Relationshsip personality and quality under consideration of preferred and existing partner characteristics] (Diploma thesis).* Hamburg, Germany: University of Hamburg.

Snyder, D. K. (1997). *Marital satisfaction inventory, revised (MSI-R).* Torrance, CA, US: Western Psychological Services.

Solomon, B. C., & Jackson, J. J. (2014). Why do personality traits predict divorce? Multiple pathways through satisfaction. *Journal of Personality and Social Psychology, 106(6)*, pp. 978-996.

Spanier, G. B. (1989). *Dyadic Adjustment Scale (DAS): Manual.* North Tonawanda, NY, US: Multi-Health Systems.

Specht, J., Schmuckle, S. C., & Egloff, B. (2011). Stability and change of personality across the life course: The impact of age and major life events on mean-level and rank-order stability of the Big Five. *Journal of Personality and Social Psychology, 101(4)*, pp. 862-882.

Swann, W. J., De la Ronde, C., & Hixon, J. G. (1994). Authenticity and positivity strivings in marriage and courtship. *Journal of personality and social psychology, 66(5)*, pp. 857-869.

Vanwinckelen, G., & Blockeel, H. (2012). On estimating model accuracy with repeated cross-validation. *BeneLearn 2012: Proceedings of the 21st Belgian-Dutch Conference on Machine Learning*, (pp. 39-44).

Vinyals, O., Toshev, A., Bengio, S., & Erhan, D. (2015). Show and tell: A neural image caption generator. *Computer Vision and Pattern Recognition (cs.CV) arXiv:1411.4555*, pp. 1-9. Retrieved from http://pages.cs.wisc.edu/~dyer/cs540/handouts/deep-learning-nature2015.pdf

Viswesvaran, C., & Ones, D. S. (1999). Meta-analyses of fakability estimates: Implications for personality measurement. *Educational and Psychological Measurement, 59(2)*, pp. 197-210.

Walther, J. B., Anderson, J. F., & Park, D. W. (1994). Interpersonal effects in computer-mediated interaction a meta-analysis of social and antisocial communication. *Communication Research, 21(4)*, pp. 460-487.

Weiss, R. L., & Cerreto, M. C. (1980). The Marital Status Inventory: Development of a measure of dissolution potential. *American Journal of Family Therapy, 8(2)*, 80-85.

Wunderlich, S. (2011). *Das Konstrukt der Beziehungs- und Bindungspersönlichkeit und sein Einfluss auf die Partnerschaftsqualität [The construct of relationship- and attachment personality and its influence on relationship quality] (Doctoral dissertation).* Hamburg, Germany: University of Hamburg.

Yarkoni, T., & Westfall, J. (2016). *Choosing prediction over explanation in psychology: Lessons from machine learning. Unpublished manuscript.* Retrieved from http://jakewestfall. org/publications/Yarkoni_Westfall_choosing_prediction.pdf

Yarkoni, T., Ashar, Y. K., & Wager, T. D. (2015). Interactions between donor Agreeableness and recipient characteristics in predicting charitable donation and positive social evaluation. *PeerJ 3:e1089*.

Youyou, W., Kosinski, M., & Stillwell, D. (2015). Computer-based personality judgments are more accurate than those made by humans. *Proceedings of the National Academy of Sciences of the United States of America, 112(4)*, pp. 1036–1040.

Zentner, M. R. (2005). Ideal mate personality concepts and compatibility in close relationships: a longitudinal analysis. *Jounal of Personality and Social Psychology, 89(2)*, pp. 242-256.

List of Figures

List of Tables

List of Abbreviations

A	Agreeableness
ACT	Activeness
ASS	Assertiveness
BEL	Believing
BB-PI	Beziehungs- und Bindungspersönlichkeitsinventar
C	Conscientiousness
CA	Cronbach's alpha
CC	Combination count
CI	Confidence interval
CV	Cross-validation
CVI	Covariance index
E	Extraversion
EMO	Emotional orientation
N-	Emotional stability (inverse Neuroticism)
FLZ	Fragebogen für Lebenszufriedenheit
GAC	Gregariousness
hh	high-high
hl	high-low
INT	Intellect
lh	low-high
ll	low-low
m	Mean/average
MAN	Manners
Max.	Maximum
Med.	Median
Min.	Minimum
MOD	Moderator
MSE	Mean squared error
ML	Machine learning
N	Number of individuals in a sample
O	Openness
PAT	Physical attractiveness

PCA	Principal component analysis
PD-I	Personality Domain Inventory
P1	Partner 1 (actor)
P2	Partner 2 (partner of actor)
RAT	rationality
RES	responsibility
SD	Standard deviation
SES	Socio-economic status
SEX	Sexuality
SIM	Similarity
SOA	Social adaptiveness
SOS	Social supportiveness
TW	Trustworthiness
T1	Time one
T2	Time two, four years later
vs.	versus
YOUME-L	YOUME List