Rank-Order Stability of Relationship Satisfaction: A Meta-Analysis of Longitudinal Studies

Janina Larissa Bühler[1,2] and Ulrich Orth[1]

[1] Department of Psychology, University of Bern

[2] Institute of Psychology, Heidelberg University

Bühler, J. L. & Orth, U. (2022). Rank-order stability of relationship satisfaction: A meta-analysis of longitudinal studies. *Journal of Personality and Social Psychology, 123*(5), 1138-1165. http://dx.doi.org/10.1037/pspp0000430

Author Note

Janina Larissa Bühler https://orcid.org/0000-0003-3684-9682

Ulrich Orth https://orcid.org/0000-0002-4795-515X

Janina Larissa Bühler is now at the Department of Psychology, Johannes Gutenberg University Mainz. This research was supported by Swiss National Science Foundation Grant P2BSP1_188102 to Janina Larissa Bühler. Data, analysis script, and research materials are available on the Open Science Framework (https://osf.io/n2z6b/).

Correspondence concerning this article should be addressed to Janina Larissa Bühler, Department of Psychology, Johannes Gutenberg University Mainz, Binger Str. 14-16, 55122 Mainz, Germany. Email: jbuehler@uni-mainz.de.

The authors thank Samantha Krauss for her valuable help with coding.

Abstract

Satisfaction with a romantic relationship often changes over time, and individuals differ in how satisfied they are in their relationship. However, no systematic review is available regarding the stability of individual differences in relationship satisfaction. Therefore, this meta-analysis synthesizes the available longitudinal data on rank-order stability of relationship satisfaction, as a function of age and relationship duration. Analyses were based on 148 samples including 153,396 participants reporting on their relationship over time. Mean age associated with the effect sizes ranged from 19 to 71 years, and mean relationship duration from 3 months to 46 years. On average, individual differences in relationship satisfaction were highly stable over time (r = .76, corrected for attenuation due to measurement error and based on an average time lag of 2.30 years). Rank-order stability varied systematically as a function of age, increasing from young to late adulthood with a slight decline during middle adulthood. Rank-order stability also varied as a function of relationship duration, increasing over the course of the relationship with a slight decline around 20 years of relationship duration. Moderator analyses suggested that relationship transitions shortly before Time 1 and sample type explained variance in rank-order stability. However, except for these two moderators, the pattern of findings was robust across all characteristics tested. In sum, this meta-analysis indicates that relationship satisfaction is a relatively stable construct, with lower stabilities in young adulthood and in the first years after beginning a relationship. This knowledge may stimulate future research on developmental processes within romantic relationships.

Keywords: relationship satisfaction; rank-order stability; longitudinal studies; meta-analysis

Rank-Order Stability of Relationship Satisfaction: A Meta-Analysis of Longitudinal Studies

For most people, romantic relationships are among the closest and most important relationships they experience in adulthood (Mund & Neyer, 2014), and satisfaction with the romantic relationship is a predictor of health, longevity, and subjective well-being (Proulx et al., 2007; Robles et al., 2014; Sbarra et al., 2011). As people go through life, their relationship satisfaction does not remain constant but typically changes over time. However, although mean-level change of relationship satisfaction has been meta-analyzed in previous research (Bühler et al., 2021), rank-order stability of relationship satisfaction has not yet been examined systematically. This is a critical limitation of the current state of knowledge because information on both indices of stability and change is needed to gain a comprehensive understanding of the development of relationship satisfaction across adulthood. Moreover, knowledge about rank-order stability of relationship satisfaction may contribute to the understanding of the nature of relationship satisfaction, by providing information about the degree to which relationship satisfaction should be conceptualized as a trait-like construct. Researchers have also debated about the most relevant time metric (i.e., age vs. relationship duration) when studying the development of relationship satisfaction (Anderson et al., 2010). Thus, it is essential to examine rank-order stability of relationship satisfaction both as a function of age and as a function of relationship duration.

The goal of the present meta-analysis was to synthesize the available longitudinal data on rank-order stability of relationship satisfaction within a given relationship to gain a robust and precise picture of the stability of individual differences in relationship satisfaction. Specifically, we sought to answer three questions: (a) What is the average rank-order stability of relationship satisfaction? (b) Does rank-order stability of relationship satisfaction vary across adulthood, as a function of age and as a function of relationship duration? (c) Does the degree of stability differ across sample and methodological characteristics?

Rank-Order Stability of Relationship Satisfaction

Two central indices of stability must be considered when examining development across adulthood: mean-level change and rank-order stability (Roberts & DelVecchio, 2000; Roberts & Nickel, 2021; Roberts et al., 2006). When applied to the concept of relationship satisfaction, indices of mean-level change reflect the average increase or decrease in relationship satisfaction in a sample of individuals over time. If mean-level change is mapped on a specific time metric (such as age or relationship duration), it is also referred to as index of normative change. In contrast, rank-order stability of relationship satisfaction reflects the stability of individual differences over time (i.e., high stability indicates that individuals in a sample tend to keep the same rank on the construct over time). Thus, if rank-order stability is high, then the relative position of individuals in the sample at a first assessment is a good predictor of the relative position at a later assessment. Hence, the indices of mean-level change and rank-order stability capture different aspects of developmental patterns in psychological constructs (Caspi & Roberts, 1999; Robins, Fraley, et al., 2001). For example, imagine three persons Heather, Tom, and Mary, of whom Heather has the highest and Mary the lowest level of relationship satisfaction. If rank-order stability is high, then Heather will still have the highest and Mary still the lowest level of relationship satisfaction at the next assessment—irrespective of the mean level of their relationship satisfaction, which might have changed for Heather, Tom, and Mary in a similar manner.

A recent meta-analysis examined the available data on mean-level change in relationship satisfaction across adulthood, focusing on the role of age and relationship duration (Bühler et al., 2021). The findings indicated that trajectories differed systematically between the time metrics. Specifically, whereas the findings showed a U-shaped trend for age, the pattern was more complex for relationship duration showing a decline in the first 10 years of a relationship, followed by an increase over the next 10 years and again a decline after 20 years of relationship duration. However, as noted above, findings on mean-level change do

not provide any information about stability and change in the relative placements of individuals (Block, 2014). For example, consider again a sample of individuals like Heather, Tom, and Mary. All three of them may decrease in their relationship satisfaction over time (i.e., the sample mean will decrease over time, indicating a lack of mean-level stability). However, if Heather, Tom and Mary decrease by the same amount, their rank ordering will remain exactly the same (indicating presence of rank-order stability). Moreover, even if the individuals decrease by slightly different amounts (i.e., indicating individual differences in change), it is possible that the rank order remains the same (indicating rank-order stability). The example illustrates that change in mean levels is theoretically independent from change in rank order.

More precisely, rank-order stability of relationship satisfaction concerns the question to which degree relationship satisfaction should be conceptualized as a trait-like construct (Fraley & Roberts, 2005). For example, intelligence is an individual-difference construct that shows particularly high rank-order stability (Neisser et al., 1996). Also, the Big Five personality traits and, to a somewhat lesser degree, self-esteem and life satisfaction are highly stable constructs that are, consequently, considered personality traits (Lucas & Donnellan, 2007; Roberts & DelVecchio, 2000; Trzesniewski et al., 2003). A construct like mood, in contrast, often changes quickly in response to the social environment (e.g., behavior of other people) and in response to intrapersonal processes (e.g., expectancies), and is therefore considered a state, not a trait. Based on a large meta-analytic dataset, Anusic and Schimmack (2016) estimated the rank-order stability of several psychological constructs (i.e., personality traits, self-esteem, life satisfaction, and affect), showing that rank-order stability decreases as the time lag increases. At the same time, their findings on the longterm rank-order stability (i.e., rank-order stability across long time intervals and corrected for measurement error) showed that the stability coefficients asymptotically approached values of .83 for personality traits, .56 for self-esteem, .52 for life satisfaction, and .42 for affect. Similar asymptotic

values have been reported in other studies on the longterm rank-order stability of these constructs (Fraley & Roberts, 2005; Fujita & Diener, 2005; Kuster & Orth, 2013; Lucas & Donnellan, 2007; Wagner et al., 2016). These findings imply that the rank-order stability of psychological constructs is lower for longer time lags, but that rank-order stability does not approach zero but nonzero values between 0 and 1, even across very long time lags (such as several decades). The differing sizes of the asymptotic values (ranging from .42 to .83) also suggest that psychological constructs differ in the degree to which they are trait-like.

There are two ways of how relationship satisfaction can be conceptualized (e.g., Fincham et al., 2018). On the one hand, relationship satisfaction is considered a construct similar to life satisfaction. Empirical data suggest that relationship satisfaction is correlated with life satisfaction at about medium size (.29 to .47; Be et al., 2013; Bühler et al., 2019). Consequently, the rank-order stability of relationship satisfaction might be similar to that of life satisfaction (Fujita & Diener, 2005; Lucas & Donnellan, 2007; Lykken & Tellegen, 1996; Schimmack & Oishi, 2005). On the other hand, relationship satisfaction strongly correlates with behavioral patterns in the relationship, such as communication styles. In fact, items of both the Marital Adjustment Test (MAT; Locke & Wallace, 1959) and the Dyadic Adjustment Scale (DAS; Spanier, 1976) correlate more strongly with communication factors than with satisfaction factors (Funk & Rogge, 2007). Given that communication patterns are less stable and often change over time (Johnson et al., 2021), rank-order stability of relationship satisfaction might also be lower and closer to rank-order stability of state-like constructs.

Moreover, when assessing the rank-order stability of a psychological construct, such as relationship satisfaction, it is essential to account for the time lag between assessments. As noted above, theory and empirical findings clearly suggest that rank-order stability is often large when the time lag is short (e.g., one year), but rank-order stability decreases as the time lag increases (e.g., Ardelt, 2000; Fraley & Roberts, 2005; Kuster & Orth, 2013; Terracciano et al., 2006). Specifically, as the time lag increases, rank-order stability typically levels off at

medium values, and this nonzero asymptote suggests that there is an enduring component of individual differences in a given construct even across very long periods. Therefore, it is crucial to account for the time lag between assessments when meta-analyzing rank-order stability of relationship satisfaction.

Mechanisms Underlying the Rank-Order Stability of Relationship Satisfaction

Three types of mechanisms are relevant for explaining the rank-order stability of a psychological construct: stochastic-contextual processes, person-environment transactions, and developmental constancy factors (Fraley & Roberts, 2005). Below, we briefly review these mechanisms and discuss how each of them applies to the rank-order stability of relationship satisfaction. It should be noted that these mechanisms are not mutually exclusive but jointly contribute to the rank-order stability of psychological constructs (Fraley & Roberts, 2005).

First, all developmental processes are influenced by stochastic-contextual processes, that is, by relatively random contextual factors, such as moving to a new place or meeting a potential mate (Lewis, 1997, 1999, 2000a, 2000b). In fact, the statistical modeling by Fraley and Roberts (2005) suggested that stochastic-contextual processes are needed to explain individual differences in a psychological construct. More precisely, when the influence of stochastic-contextual processes is ignored, individual differences in a construct would be perfectly stable over time. Hence, stability and change in a psychological construct depend on the stability of the context, and the degree of change and stability depends on how stable the environmental conditions are.

Second, rank-order stability of psychological constructs also depends on person-environment transactions, which means that individuals actively shape their environmental conditions and that, simultaneously, these environmental conditions affect the individual (Caspi & Bem, 1990; Caspi et al., 1989; Neyer & Asendorpf, 2001; Neyer et al., 2014). These dynamic transactions between individuals and their environment (e.g., relationships) foster

consistency in individual differences. There are at least two transactive processes that are relevant for explaining the rank-order stability of relationship satisfaction (Caspi & Bem, 1990). First, the way how individuals perceive their romantic relationship is influenced by idiosyncratic social-cognitive biases (a transaction called reactive process). For example, individuals who are satisfied with their relationship tend to perceive their partner and their relationship through rose-colored glasses, which increases the likelihood of more positive relationship experiences in the future. In contrast, individuals who are unhappy with their relationship, show a negative bias in the perception of their relationship, which may cause disappointment and relationship conflicts. In both cases, the person makes relationship experiences that are congruent with their prior relationship satisfaction, which contributes to the stability of individual differences in relationship satisfaction (e.g., Ickes et al., 1997; Swann & Read, 1981). Second, individuals actively select themselves into environments, including romantic relationships (a transaction called proactive process). For example, individuals who believe that they are lovable and that others are trustworthy, tend to select more supporting and trustworthy partners, which leads to more positive relationship experiences (Erol & Orth, 2016; Sandra L. Murray et al., 2000). In contrast, individuals who have more negative views of the self and others, will select untrustworthy partners, which will cause more negative relationship experiences. In both cases, individuals tend to make relationship experiences that are consistent with their pre-existing beliefs, which again contributes to the stability of individual differences in relationship satisfaction. More generally, the theoretical and statistical model by Fraley and Roberts (2005) suggested that person-environment transactions amplify the degree of rank-order stability of psychological constructs.

Third, rank-order stability is also influenced by developmental constancy factors (e.g., genetic predispositions and early formative experiences), which emphasize the role of latent resiliency and vulnerability factors (Bowlby, 1973; McGue et al., 1993; Roberts & Caspi,

2003; Roberts & Wood, 2006). More precisely, developmental constancies may predispose people to perceive themselves and their environments (e.g., their relationship and their relationship partners) in a specific manner and to behave in ways that influence the quality of their romantic relationships. For example, individuals who are high in neuroticism tend to negatively interpret ambiguous relationship situations (Finn et al., 2013), which contributes to lower relationship satisfaction. Such vulnerability and latent resiliency factors are a constant influence on people's perceptions and behavior in the relationship domain (McNulty, 2016), which contributes to the stability of individual differences in relationship satisfaction. The analyses by Fraley and Roberts (2005) suggested that developmental constancy factors are needed to explain the typical pattern of rank-order stability of psychological constructs. In fact, when constancy factors were omitted from their model, rank-order stability quickly approached zero as the interval between assessments became longer.

Applied to romantic relationships, this reasoning suggests that the rank-order stability of relationship satisfaction depends on the stability of all factors that influence the quality of relationships, that is, (a) individual characteristics of the two partners, (b) characteristics of the relationship, and (c) contextual factors outside of the relationship. For example, the relationship science literature suggests that individual characteristics of the partners such as emotional stability, conscientiousness, agreeableness, self-esteem, and attachment security (i.e., low levels of both attachment-related anxiety and attachment-related avoidance) significantly influence the quality of their relationship and, consequently, the partners' satisfaction with their relationship (e.g., Erol & Orth, 2016; Li & Chan, 2012; McNulty, 2016; Weidmann, Ledermann, et al., 2017). If individual differences in these factors are quite stable over time (as suggested by empirical research; e.g., Fraley & Roberts, 2005; Kuster & Orth, 2013; Roberts & DelVecchio, 2000; Scharfe & Bartholomew, 1994), this suggests that individual differences in relationship satisfaction will likewise be relatively stable. Similarly, research suggests that individual characteristics, including emotional stability,

conscientiousness, agreeableness, self-esteem, and attachment security, contribute to characteristics of the relationship (i.e., the pattern of relationship behavior that has evolved in a specific relationship), such as patterns of communication, coping styles, responsiveness, and conflict resolution (e.g., Campbell et al., 2005; S. L. Murray et al., 2000; Vater & Schröder-Abé, 2015). Relationship patterns, in turn, influence people's relationship satisfaction (e.g., Debrot et al., 2012; Gottman & Levenson, 1999; Vater & Schröder-Abé, 2015). Hence, because individual differences in personality characteristics are relatively stable, patterns of relationship behavior might also be relatively stable over time, which would further contribute to rank-order stability of relationship satisfaction. In contrast, theory suggests that contextual factors outside of the relationship may destabilize the relationship (Bodenmann, 1995; Hill, 1958; Randall & Bodenmann, 2009). For instance, the vulnerability-stress-adaptation model (Karney & Bradbury, 1995; for a recent extension, see McNulty et al., 2021) emphasizes that stressful life events (e.g., birth of a child) may impair adaptive processes within the relationship (e.g., coping styles), which may compromise the relationship satisfaction of couple members. Thus, samples that experienced potentially stressful changes in the relationship context (e.g., samples of couples who had their first baby) might show lower rank-order stability of relationship satisfaction than samples who did not experience significant changes in the relationship context.

Does Rank-Order Stability of Relationship Satisfaction Vary Across Adulthood?

In addition to estimating the average rank-order stability of relationship satisfaction, it is essential to understand how stability varies across adulthood. As noted above, in this meta-analysis we examined rank-order stability of relationship satisfaction within a given relationship as a function of age and relationship duration. Clearly, both time metrics are strongly correlated: People of higher age, compared to people of younger age, are often in relationships of longer duration, simply because they are older. Nevertheless, people separate from their partner and begin a new romantic relationship across the entire period of adulthood

(Carr & Utz, 2020; Mehta et al., 2020). Moreover, while some of the mechanisms underlying the rank-order stability of relationship satisfaction refer to age, others refer to relationship duration. Therefore, it is important to separate the two time metrics, both conceptually and empirically.

Rank-Order Stability of Relationship Satisfaction as a Function of Age

The developmental literature allows to derive hypotheses about how rank-order stability of relationship satisfaction varies as a function of age. For instance, developmental task theory posits that each life stage entails new developmental demands and societal expectations (Erikson, 1968; Havighurst, 1972; see also Hutteman et al., 2014). Specifically, a key developmental task in young adulthood (i.e., age 18 to 40 years) is to establish long-lasting social relationships, including a committed romantic relationship (Ebner et al., 2006; Heckhausen et al., 1989). At the same time, young adulthood is also characterized by exploring different life paths and options, which sometimes implies leaving and entering romantic relationships more readily (Arnett, 2000; Halpern-Meekin et al., 2013; Shulman & Connolly, 2013).

In contrast, middle adulthood (i.e., age 40 to 65 years) involves the developmental tasks of generativity and consolidation, expressed in caring for the next generation and maintaining satisfactory social relationships, including marriage or a marriage-like relationship (Erikson, 1968; McAdams, 2015). Moreover, in middle adulthood, individuals usually develop an executive personality, which is, among other aspects defined by an increase in mastery, competence, and control (Neugarten, 1968). This, in turn, increases the capacity to handle multiple pressures and to cope successfully with difficult personal and interpersonal experiences (Roberts & DelVecchio, 2000). Yet, middle adulthood is also a time of potential crisis (Freund & Ritter, 2009; Levinson et al., 1976), resulting from the many responsibilities in family, work, and community contexts, which may lead to stress, conflict, and instability (Freund & Nikitin, 2012).

Finally, late adulthood (i.e., age 65 years and older) is often characterized by a greater salience of loss-related issues (such as loss of beloved ones) and the perception of limited remaining time (Carstensen et al., 1999; Charles & Carstensen, 2010). Therefore, key developmental tasks in late adulthood involve the avoidance of, and adjustment to, losses and the selective investment of time and energy into life domains (Ebner et al., 2006; Freund, 2008; Heckhausen et al., 1989; Ogilvie et al., 2001). As a result, older adults tend to focus more strongly on present-oriented, rather than future-oriented, goals and invest more time and energy in positive relationships with close others than in social interaction with acquaintances (Carstensen et al., 1999; Frederick et al., 2017; Fung et al., 1999).

Taken together, the typical life situations and developmental tasks in young, middle, and late adulthood suggest that individuals invest increasingly in establishing and maintaining a romantic relationship as they go through life. This, in turn, should lead to more stability in people's relationship conditions. Given that romantic relationships differ substantially with regard to their relationship quality (i.e., some will be fulfilling and satisfying for the partners, whereas others will involve some level of conflict and be less satisfying), the developmental trend towards consolidating and maintaining romantic relationships suggests that individual differences in relationship satisfaction become more stable with age.

Rank-Order Stability of Relationship Satisfaction as a Function of Relationship Duration

Perspectives from relationship science allow to derive hypotheses on how rank-order stability of relationship satisfaction varies as a function of relationship duration. Specifically, the gradual disillusionment model (Huston et al., 2001; Huston & Houts, 1998) suggests that baseline levels of relationship satisfaction decrease over the course of the relationship (Diekmann & Mitter, 1984; Kurdek, 1998, 1999). The strongest decline often occurs over the first 10 years of a relationship and, consequently, risk of separation peaks at around 10 years after beginning a relationship (i.e., roughly corresponding to 7 years of marriage duration, given that couples ususally have been together a few years before marrying; Kulu, 2014). This

implies that relationships of longer duration (10 years and more) predominantly represent the "surviving" and more satisfied couples who have created a more stable relationship context over time. As reviewed above, Fraley and Roberts' (2005) model suggests that stable environments contribute to higher rank-order stability of individual-difference constructs. Therefore, we expected an increase in rank-order stability of relationship satisfaction over the course of the relationship.

Moderators of Rank-Order Stability of Relationship Satisfaction

As noted above, individual factors and characteristics of the romantic relationship might explain why rank-order stability of relationship satisfaction differs as a function of age and relationship duration. Therefore, in this meta-analysis we also tested whether sample and methodological characteristics are moderators of rank-order stability. For some of the moderators, the literature allows to derive hypotheses about the significance and direction of effects (i.e., living arrangement, marital status, presence of children, occurrence of relationship transitions, and type of measure). For other moderators, however, no hypotheses could be derived (e.g., ethnicity or gender). Nevertheless, to gain information about the robustness and generalizability of the findings, we tested the full set of moderators.

Living Arrangements, Marital Status, and Presence of Children

People's living arrangement, their marital status, and presence of children are sample characteristics that may contribute to the explanation of rank-order stability of relationship satisfaction. Specifically, couples who live in the same household, are married, and/or have children might live in more stable relationship environments compared to couples who live in separate households, are unmarried, and/or do not have children. Moreover, couples who live in the same household, are married, and/or have children might encounter greater legal, financial, and social barriers to separation, which may prevent relationship break-up even if they are unsatisfied in their relationship (Rusbult, 1980, 1983). Together, these factors may contribute to more stable individual differences in relationship satisfaction.

Occurrence of Relationship Transitions

A sample characteristic that may contribute to lower rank-order stability of relationship satisfaction is the occurrence of relationship transitions. Relationship transitions generate new demands for individuals and couples, such as changes in family and work roles, including new arrangements for household and work duties (Sanchez & Thomson, 1997). Also, new parents often undergo stages of elevated stress and conflict (Doss et al., 2009). These demands likely challenge the stable and consistent environment of the relationship (Belsky & Rovine, 1990; Cast, 2004). Thus, the occurrence of relationship transitions may destabilize the system of transactions between the person and the environment (Fraley & Roberts, 2005).

Type of Measure

A methodological characteristic that may moderate the findings is the type of measure. In general, measures of relationship satisfaction can be grouped into ad-hoc measures and established measures. Established measures can be further classified into global satisfaction measures, such as the Relationship Assessment Scale (RAS; Hendrick, 1988), and adjustment measures, such as the DAS (Spanier, 1976). Global satisfaction measures rely on an intrapersonal approach, reflecting people's subjective evaluations of the relationship in general. Adjustment measures, on the other hand, rely on an interpersonal or relationship approach, reflecting typical patterns of interactions in the relationship, such as communication and conflict styles (Fincham et al., 2018). Global satisfaction measures likely reflect the more trait-like aspect of relationship satisfaction, while adjustment measures reflect the more state-like aspect. Consequently, the use of global satisfaction measures should lead to greater estimates of rank-order stability.

Sample Type, Ethnicity, Gender, and Baseline Mean of Relationship Satisfaction

We also tested for the moderating effects of sample type (i.e., nationally representative vs. nonrepresentative), ethnicity, gender, and baseline mean of relationship satisfaction.

Although we did not have directional hypotheses about the findings for these variables, testing these variables provides important information about the robustness and generalizability of the findings (see also Bühler et al., 2021; Orth et al., 2018). Specifically, representative samples typically allow for more valid conclusions compared to nonrepresentative samples, such as community samples and samples of college students (Orth et al., 2018). Ethnicity and gender are key demographic characteristics that are of interest to many researchers, so it is important to test whether meta-analytic findings differ between ethnic groups and between women and men. For example, research suggests that dynamic processes in romantic relationships vary by ethnicity (e.g., Orengo-Aguayo, 2015; see also Karney & Bradbury, 2020). Finally, testing whether the meta-analytic findings hold across the general level of relationship satisfaction in a sample (as indicated by the baseline mean) is important, because some research suggests that baseline means in relationship satisfaction may influence the dynamic processes in the relationship (e.g., Lavner et al., 2012). Thus, testing these sample characteristics as moderators provides important information about the generalizability of the findings.

The Present Research

The goal of this research was to synthesize the available longitudinal data on rank-order stability of relationship satisfaction in adulthood. In the analyses, we will examine how rank-order stability varies as a function of age and relationship duration. As noted above, when estimating rank-order stability, it is essential to consider the time lag between assessments. Therefore, we will conduct two sets of effect size analyses: without versus with controlling for time lag (for similar procedures, see Roberts & DelVecchio, 2000; Trzesniewski et al., 2003). The analyses that control for time lag will yield estimates of rank-order stability as if all samples had the same time lag between assessments (by centering time lag at the mean across effect sizes, i.e., 2.30 years). Finally, moderator analyses will provide information about the robustness of the findings, by testing whether rank-order stability

differs across sample and methodological characteristics. It is important to note that some potentially relevant moderators could not be examined in this meta-analysis, such as personality variables, socioeconomic status, and sexual orientation (Chen & van Ours, 2018; Conger et al., 2010; Karney & Bradbury, 1995). The reason is that (a) information on the characteristics was not reported in most primary studies (i.e., personality variables), (b) the information that was available was not comparable across most primary studies (i.e., socioeconomic status), or (c) the very low number of samples that provided information on the characteristic would not have allowed for reliable conclusions (i.e., sexual orientation).

The present meta-analysis advances research on romantic relationships in several ways. Although rank-order stability is a sample (or population) characteristic, knowledge about how rank-order stability of relationship satisfaction changes as a function of age and relationship duration, and about which individual and environmental factors moderate rank-order stability, contributes to understanding the development of relationship satisfaction and, more generally, relationships. For example, if the stability of individual differences is particularly low in a specific developmental period (e.g., young adulthood), then this suggests that the individual trajectories are more variable, and likely more malleable, compared to developmental periods in which individual differences are very stable. Similarly, if the stability of individual differences is relatively low in a specific relationship situation (e.g., in the period after relationship transitions such as the transition to parenthood), then this suggests that relationship interventions might be more impactful in this situation compared to other situations in which rank-order stability is high. Moreover, this meta-analysis will allow to evaluate whether rank-order stability of relationship satisfaction is of about similar size as the rank-order stability of personality characteristics (such as the Big Five) or whether individual differences in relationship satisfaction are less stable over time. If individual differences in relationship satisfaction are as stable as individual differences in personality

characteristics, then this would suggest that relationship satisfaction within a given relationship should be considered a trait-like construct.

Method

This meta-analysis used anonymized data and was therefore exempt from receiving approval by the Ethics Committee of the authors' institution (Faculty of Human Sciences, University of Bern), in accordance with national law.

Transparency and Openness

We follow the Journal Article Reporting Standards (Appelbaum et al., 2018; Kazak, 2018) and describe how we obtained the samples included in the present meta-analysis. Data, analysis script, and research materials (e.g., coding manual, information on study variables) are available on the Open Science Framework (OSF, https://osf.io/n2z6b/). The design and analyses of the present research were not pre-registered. Data were analyzed using R (R Development Core Team, 2020), and the meta-analytic computations were conducted with the metafor package (Viechtbauer, 2010).

The meta-analytic data set is based partially on data from another meta-analysis of longitudinal studies on relationship satisfaction. Specifically, we used data on sample and methodological characteristics that were also used in a meta-analysis on means and mean-level change in relationship satisfaction (Bühler et al., 2021). For the present research, the meta-analytic data set has been extended by including further studies that met the inclusion criteria of the present meta-analysis and by coding the information required for meta-analyzing rank-order stability of relationship satisfaction. With regard to effect size data and analyses, there is no overlap between the present research and Bühler et al. (2021). For reasons of completeness and clarity, we provide all relevant methodological information below, even if information on some of the search and coding procedures is also reported in Bühler et al. (2021).

Search and Selection Procedure

The flow diagram shown in Figure 1 summarizes the four steps of the search and selection procedure in the present meta-analysis: identification, screening, eligibility, and inclusion of studies.

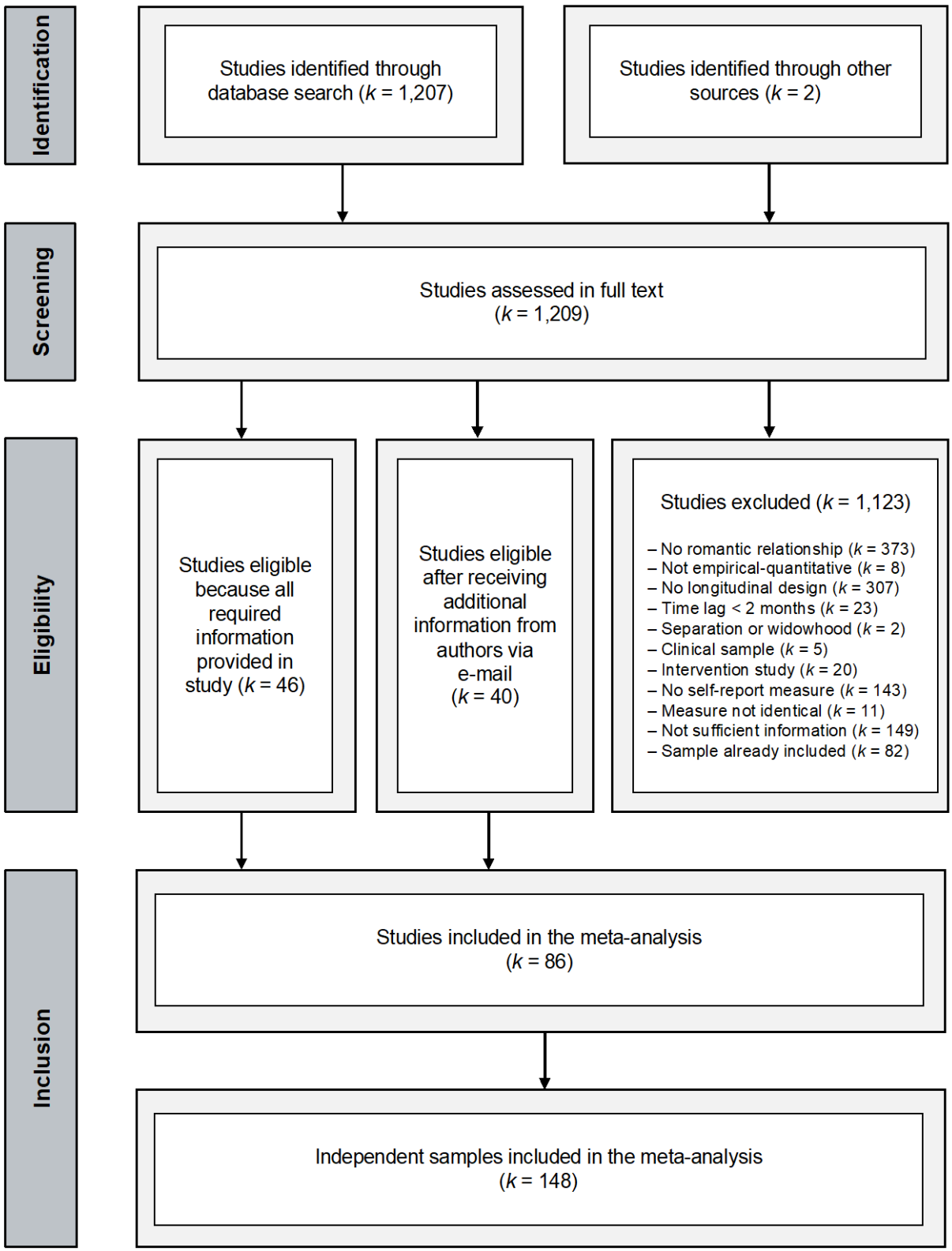
Identification of Studies

We searched for English-language journal articles, books, book chapters, and dissertations in the database PsycINFO to find relevant studies. The search was conducted on September 9, 2019, and the following search terms were used: *relationship satisfaction, marital satisfaction, relationship quality, marital quality, dyadic adjustment, marital adjustment,* and *marital relations.* The search was restricted to empirical-quantitative and longitudinal studies with non-clinical samples, by using the limitation options *empirical study*, *quantitative study*, *longitudinal study*, and *non-disordered population* in PsycINFO. The search yielded 1,207 potentially relevant studies, and two additional potentially relevant studies were identified through other sources. Thus, the final data set consisted of 1,209 potentially relevant studies, including 53 dissertations.[1]

[1] Among the potentially relevant studies, the oldest study had been published in 1966. However, none of the studies published before 2002 were included in the meta-analysis, for the following reasons. First, most of the studies did not meet the inclusion criteria (e.g., because the studies were not longitudinal). Second, some studies would have met the inclusion criteria, but the information on effect sizes provided in the study was insufficient and although we contacted the authors of the studies, we did not receive the required information (e.g., because the authors were no longer in academia). Third, some of the studies would have met the inclusion criteria but a more recent study (i.e., published 2002 or later) using the same data was included because the more recent study reported more complete information on sample characteristics and effect sizes.

Figure 1

Flow Diagram of the Search and Selection Procedure



Note. The diagram has been adapted from Moher et al. (2009).

Screening and Eligibility of Studies

To decide on the eligibility of the studies, all studies were assessed in full text by the first author or a second rater based on the inclusion criteria described below.[2] To decide whether a study met the criteria for being included in the meta-analytic data set, the raters followed standardized procedures. Studies were included if the following twelve criteria were fulfilled: First, participants reported on a romantic relationship (i.e., measures of satisfaction with non-romantic relationships, such as parent-child relationships, were not of interest). Second, the study was empirical-quantitative. Third, the study used a longitudinal design (i.e., two or more assessments of the same sample). Fourth, the time lag between assessments was 2 months or more. More precisely, for the meta-analysis we used data from assessments that were separated by at least 2 months; these assessments were coded as Time 1, Time 2, Time 3, etc. Fifth, the sample did not, as a whole, experience separation or widowhood. Sixth, the sample was not clinical. Seventh, the study was not an intervention study. Eighth, relationship satisfaction was assessed by self-report. Ninth, the measure of relationship satisfaction was identical across assessments. Tenth, sufficient information was given to compute the effect size. Eleventh, information on effect size data was consistent throughout the study. Twelfth, the sample was not already included in the meta-analytic data set (specifically, when a sample was used in more than one study, we selected the study that provided information on the largest sample size or, if identical, the most comprehensive information on sample and effect size data).

To obtain estimates on interrater agreement on eligibility, a random sample of 60 studies were rated by both raters, suggesting high interrater agreement on inclusion versus exclusion of articles in the meta-analytic data set (i.e., 59 of 60 articles, resulting in $\kappa = .92$). The diverging assessment was discussed until consensus was reached.

[2] At the time of coding (i.e., October 2019 to February 2020), the qualifications of the raters were as follows: The first author had a Ph.D. in Psychology and the second rater had a Master's degree in Psychology.

Inclusion of Studies

Of the potentially relevant studies, 46 studies could be included immediately because they provided sufficient information on the effect size (i.e., correlation coefficient between at least two assessments) and information on mean age and/or mean relationship duration. In the case of studies that met the criteria except for providing information on the effect size, mean age, or mean relationship duration, we contacted the authors of the study with a request for the missing information (if sufficient contact information was available in the study or could be found elsewhere, e.g., on the website of the authors' university). This procedure led to the inclusion of 40 additional studies. In sum, the search procedures resulted in a total of 86 eligible studies, including 146 independent samples.

Coding Procedure

The 146 samples were coded by the first author or the second rater. The following data were coded: year of publication, publication type, sample size, sample type, country, ethnicity, sexual orientation, proportion of female participants, proportion of participants living together with their partner in the same household, proportion of married participants, proportion of participants with children, occurrence of a relationship transition (i.e., marriage or birth of a child) between Time 1 and any of the following assessments (referred to as *transition*), occurrence of a relationship transition (i.e., marriage or birth of a child) shortly before Time 1 (referred to as *post transition*), type of relationship transition, time lag between assessments, dyadic nature of sample, measure of relationship satisfaction, reliability coefficient of relationship satisfaction averaged across assessments, range of scale (i.e., the scale's minimum possible score and the scale's maximum possible score), mean of relationship satisfaction at Time 1, mean age of participants at Time 1, mean relationship duration at Time 1, year of Time 1 assessment, and correlation coefficient indicating the rank-order stability of relationship satisfaction between assessments.

In the case of dyadic female-male samples, we coded the data for female and male participants separately to increase the power of moderator analyses testing for gender differences in the findings. In all other cases of subsamples (e.g., young adults and middle-aged adults), we coded the full sample.[3] If information on year of Time 1 assessment was not reported, we estimated these data as follows: Year of Time 1 assessment = publication year – 3 years – time lag between the first and last measurement occasion of the study (based on the assumption that studies are, on average, published 3 years after data collection has been completed; for a similar procedure, see Orth et al., 2018). Moreover, for the effect size analyses we needed participants' mean age and mean relationship duration at the initial assessment of each effect size interval (e.g., mean age and mean relationship duration at Time 4 for the effect size interval from Time 4 to Time 5). We used the information on mean age and mean relationship duration at Time 1 and the information on time lags between assessments to compute these values.

To obtain interrater agreement in this step of coding, a random sample of 40 studies was rated by both coders, suggesting high interrater agreement, with $\kappa = 1.00$ for categorical variables (except for one variable, see below) and $r \geq .99$ for continuous variables. For sample type, interrater agreement was $\kappa = .90$, resulting from one diverging assessment (one coding was "community sample," whereas the other was "college/university students"). All diverging assessments were discussed until consensus was reached.

As described in Bühler et al. (2021), we used the following strategies to obtain data on mean relationship duration if these data were missing in the study. First, we contacted the authors of the study with a request for the information. This resulted in data on relationship duration for 18 additional samples. Second, many studies provided information on proxies for relationship duration, that is, duration of living together (15 samples) and/or marriage

[3] The only exception was one study, in which we coded Israeli and German couples separately because the time lag between assessments differed for the subsamples.

duration (42 samples). On the basis of published data on how relationship duration, duration of living together, and marriage duration are related, we used these proxies to obtain estimates of relationship duration. More precisely, nationally representative data from Germany (Schröder & Schmiedeberg, 2015; see also Heuveline & Timberlake, 2004) provide estimates about the average time between beginning a relationship, moving in together, and marrying: On average, people are in a relationship for 1.25 years before they move in together, and they are in a relationship for 3.8 years before they marry. We used these estimates for creating an overall relationship-duration variable. That is, if information on relationship duration was missing, but information on duration of living together was available, we estimated relationship duration by adding 1.25 years to the value of duration of living together, and if information on marriage duration was available, we estimated relationship duration by adding 3.8 years to the value of marriage duration.[4] After using these procedures, information on relationship duration was available for 106 samples.

Effect Size Measure

As effect size measure, we used the correlation between two assessments of relationship satisfaction and included all available correlation coefficients that were reported for the sample. However, if measures are not perfectly reliable, then the observed correlation underestimates the true correlation. Because we were interested in estimates of the true rank-order stability of relationship satisfaction, we corrected the correlations for attenuation resulting from unreliability of the measures (Hunter & Schmidt, 1990, 2004, 2014). To obtain the most accurate estimate of reliability of a given measure, we used the following strategies. If available, we used the reliability coefficient of the measure as reported in the study (k =

[4] As described, this procedure was based on estimates about the average time between beginning a relationship, moving in together, and marrying. Therefore, we conducted sensitivity analyses by using 2 and 6 years (instead of 1.25 and 3.8 years) as estimates of the average difference between relationship duration and duration of living together, and relationship duration and marriage duration, respectively. The mean of the relationship-duration variable used in the sensitivity analyses was 13.59 years (SD = 9.71, range = 0.26–48.40). The results of the sensitivity analyses are reported in Table S1 and showed that the pattern of findings was very similar to the findings from the main analyses.

134). If no reliability coefficient was provided in the study, we used one of the following two methods: (a) When an established scale was employed (k = 4), we used the average reliability coefficient of this measure as reported in other studies in the meta-analytic data set using the same measure. (b) When a single item ad-hoc measure was employed (k = 10), we used .75 as reliability estimate because the literature suggests that single item measures of constructs that are highly schematized—i.e., constructs that can readily be reported by lay people, such as relationship satisfaction—often have a reliability in the range of .70 to .80 (Lucas & Donnellan, 2007; Robins, Hendin, et al., 2001; Woods & Hampson, 2005; for a similar procedure, see Orth, 2018).

The disattenuated correlation coefficient is given by

$$r'_i = \frac{r_i}{rel_i},$$

where r'_i is the disattenuated correlation coefficient in study i, r_i is the observed correlation coefficient in study i, and rel_i is the averaged reliability coefficient of the measure in study i (Lipsey & Wilson, 2001). For the meta-analytic computations, the disattenuated correlation coefficients were transformed to Fisher's Z_r values (Fisher, 1921; Hedges & Olkin, 1985). After the meta-analytic computations, the effect size estimates were converted back to the correlation metric.

Meta-Analytic Procedure

As noted above, for many samples the data set included more than one correlation coefficient of relationship satisfaction, which yielded a multilevel data structure (i.e., effect sizes nested in samples). To account for the multilevel structure in the meta-analytic computations, we used the "rma.mv" function in the metafor package (Viechtbauer, 2010). Following Lipsey and Wilson (2001), we used multilevel random-effects models to estimate weighted mean effect sizes and multilevel mixed-effects models to test for moderators.

When meta-analyzing correlation coefficients in the metric of Fisher's Z_r values, the within-study variance is given by

$$v_i = \frac{1}{n_i - 3},$$

where n_i is the sample size of study i. However, when meta-analyzing correlation coefficients that have been corrected for attenuation, the within-study variance is given by

$$v'_i = \frac{v_i}{rel_i^2},$$

where v'_i is the disattenuated within-study variance for study i, v_i is the attenuated within-study variance for study i (as given above), and rel_i is the averaged reliability coefficient of the measure in study i (Lipsey & Wilson, 2001).

Results

Description of Studies

The meta-analytic data set included 148 samples, drawn from 86 studies and providing 402 effect sizes. In sum, the samples included 153,396 participants, and sample sizes ranged from 32 to 84,711 (M = 1,050, SD = 7,001, Mdn = 179). An overview of the samples is given in Table 1, and all effect sizes are reported at OSF (https://osf.io/n2z6b/).

Table 1

Descriptive Information for the Samples Included in the Meta-Analysis

Study	Sample size	Mean age T1	Mean relationship duration T1	Female (in %)	Sample type	Country	Ethnicity	Measure	Reliability
Andres (2014)	153	34.00	12.00	100	Community	NLD	White	ENRICH	.80
Be et al. (2013), female	1,385	63.20	n.a.	100	National	GBR	n.a.	Ad-hoc	.80
Be et al. (2013), male	1,385	65.70	n.a.	0	National	GBR	n.a.	Ad-hoc	.71
Bikos et al. (2007)	32	38.63	n.a.	100	Community	TUR	White	KMSS	.93
Bloch et al. (2014), female	156	52.57	34.51	100	Community	USA	White	MAT/MRI	.85
Bloch et al. (2014), male	156	52.57	34.51	0	Community	USA	White	MAT/MRI	.80
Blumenstock and Papp (2017), female	373	24.30	4.30	100	Community	USA	Other	Ad-hoc	.88
Blumenstock and Papp (2017), male	373	26.50	4.30	0	Community	USA	Other	Ad-hoc	.85
Bodi et al. (2010)	389	35.04	10.87	100	Community	n.a.	White	RAS	.88
Bouchard (2014), female	151	28.00	6.25	100	Community	CAN	White	DAS	.91
Bouchard (2014), male	151	31.00	6.25	0	Community	CAN	White	DAS	.89
Bouchard et al. (2006), female	119	28.18	7.25	100	Community	CAN	n.a.	DAS	.78
Bouchard et al. (2006), male	119	30.40	7.25	0	Community	CAN	n.a.	DAS	.73
Bower et al. (2013), female	97	n.a.	7.81	100	Community	USA	White	DAS	.87
Bower et al. (2013), male	99	n.a.	7.81	0	Community	USA	White	DAS	.88
Brown et al. (2019)	88	54.98	n.a.	63	Community	USA	White	CSI	.84
Busby and Gardner (2008), female	275	26.29	4.25	100	Community	USA	White	REQ	.88
Busby and Gardner (2008), male	275	28.32	4.25	0	Community	USA	White	REQ	.88
Buyukcan-Tetik et al. (2017), female	195	29.97	6.51	100	Community	NLD	White	DAS	.87
Buyukcan-Tetik et al. (2017), male	195	32.91	6.51	0	Community	NLD	White	DAS	.87
Byers (2005)	87	37.70	12.20	62	Community	n.a.	White	GMRS	.96

Choi (2016)	2,078	31.35	7.45	100	Community	KOR	Asian	KMSS-R	.78
Christopher et al. (2015), female	96	29.34	3.52	100	Community	USA	White	MOQ	.92
Christopher et al. (2015), male	96	31.23	3.52	0	Community	USA	White	MOQ	.92
Crocker et al. (2017), Study 1	132	19.45	1.53	76	Community	USA	White	QMI	.96
DeMaris (2010)	704	35.63	16.86	62	National	USA	White	Ad-hoc	.87
Doohan et al. (2010), female	102	38.76	13.00	100	Community	USA	Other	MAT	.73
Doohan et al. (2010), male	102	41.20	13.00	0	Community	USA	Other	MAT	.74
Erol and Orth (2014), Study 2, female	6,115	40.30	15.90	100	National	USA	White	Ad-hoc	.75
Erol and Orth (2014), Study 2, male	6,115	43.00	15.90	0	National	USA	White	Ad-hoc	.75
Fallis et al. (2016), female	113	35.73	10.47	100	Community	CAN	White	QMI	.95
Fallis et al. (2016), male	113	37.96	10.47	0	Community	CAN	White	QMI	.95
Fincham and Beach (2007), female	84	41.10	n.a.	100	Community	USA	White	MAT	.90
Fincham and Beach (2007), male	84	43.30	n.a.	0	Community	USA	White	MAT	.90
Gao and Cummings (2019), female	237	37.82	16.80	100	Community	USA	White	MAT	.75
Gao and Cummings (2019), male	237	40.15	16.80	0	Community	USA	White	MAT	.75
Girme et al. (2014), Study 2, female	66	22.25	2.83	100	Community	NZL	n.a.	PRQC	.81
Girme et al. (2014), Study 2, male	66	22.25	2.83	0	Community	NZL	n.a.	PRQC	.78
Gray and Ozer (2019), female	325	29.60	n.a.	100	Community	USA	Other	CSI	.91
Gray and Ozer (2019), male	325	31.40	n.a.	0	Community	USA	Other	CSI	.91
Greving Mehall et al. (2009), female	157	30.59	9.37	100	Community	USA	White	MAT	.78
Greving Mehall et al. (2009), male	157	32.58	9.37	0	Community	USA	White	MAT	.72
Gustavson et al. (2016), female	238	46.00	n.a.	100	Community	NOR	White	RAS	.90
Gustavson et al. (2016), male	194	48.00	n.a.	0	Community	NOR	White	RAS	.90
Hagemeyer et al. (2013), female	547	39.40	11.40	100	Community	DEU	White	Ad-hoc	.75
Hagemeyer et al. (2013), male	547	41.60	11.40	0	Community	DEU	White	Ad-hoc	.75
Hakanen et al. (2011)	1,632	44.90	n.a.	72	Community	FIN	White	RSI	.96
Hammond and Overall (2014), female	88	21.08	2.58	100	Community	NZL	n.a.	PRQC	.88

Hammond and Overall (2014), male	88	22.73	2.58	0	Community	NZL	n.a.	PRQC	.88
Hernandez-Kane and Mahoney (2018)	67	29.85	5.05	67	Community	USA	White	KMSS	.97
Hsiao (2017)	614	42.83	n.a.	82	Community	TWN	Asian	Ad-hoc	.75
Impett et al. (2012), Partner 1	80	23.89	n.a.	n.a.	Community	CAN	Other	RSI	.90
Impett et al. (2012), Partner 2	76	23.79	n.a.	n.a.	Community	CAN	Other	RSI	.90
Ivanova (2016)	4,116	46.90	23.83	60	National	NLD	White	Ad-hoc	.95
Jayamaha and Overall (2015), Study 1	156	22.21	2.30	65	Student	NZL	n.a.	PRQC	.84
Jayamaha and Overall (2015), Study 2, female	174	22.43	3.00	100	Community	NZL	n.a.	PRQC	.85
Jayamaha and Overall (2015), Study 2, male	174	23.82	3.00	0	Community	NZL	n.a.	PRQC	.85
Jenkins et al. (2020), female	168	40.75	15.41	100	Community	USA	Black	CRDQ	.91
Jenkins et al. (2020), male	168	43.57	15.41	0	Community	USA	Black	CRDQ	.91
Jensen and Rauer (2015a), female	64	70.00	46.20	100	Community	USA	White	MSQO	.82
Jensen and Rauer (2015a), male	64	71.00	46.20	0	Community	USA	White	MSQO	.92
Jensen and Rauer (2015b)	67	20.80	3.03	100	Student	USA	White	IRQ	.94
Johnson and Anderson (2013), female	610	28.45	2.58	100	Community	USA	White	Ad-hoc	.90
Johnson and Anderson (2013), male	610	30.52	2.58	0	Community	USA	White	Ad-hoc	.90
Kanat-Maymon et al. (2016), Israeli, female	102	41.53	n.a.	100	Community	ISR	White	ENRICH	.79
Kanat-Maymon et al. (2016), Israeli, male	103	43.11	n.a.	0	Community	ISR	White	ENRICH	.79
Kanat-Maymon et al. (2016), German, female	209	39.16	n.a.	100	Community	DEU	White	ENRICH	.79
Kanat-Maymon et al. (2016), German, male	210	41.14	n.a.	0	Community	DEU	White	ENRICH	.79
Kerkhof et al. (2011), female	199	29.20	5.77	100	Community	NLD	White	DAS	.84
Kerkhof et al. (2011), male	199	32.07	5.77	0	Community	NLD	White	DAS	.86
Kluwer and Johnson (2007), female	262	28.80	6.46	100	Community	NLD	White	RSI	.89
Kluwer and Johnson (2007), male	262	31.20	6.46	0	Community	NLD	White	RSI	.89
Kouros (2011), female	296	37.84	14.38	100	Community	USA	White	MAT	.78
Kouros (2011), male	296	40.22	14.38	0	Community	USA	White	MAT	.78
Lavner and Bradbury (2010), female	232	25.50	4.13	100	Community	USA	Other	MAT	.78

Lavner and Bradbury (2010), male	232	27.00	4.13	0	Community	USA	Other	MAT	.78
LeBaron et al. (2014)	67	46.00	27.70	100	Community	USA	White	Ad-hoc	.87
Li et al. (2018), female	268	28.08	4.93	100	Community	CHN	Asian	QMI	.96
Li et al. (2018), male	268	29.59	4.93	0	Community	CHN	Asian	QMI	.94
Lickenbrock and Braungart-Rieker (2015), female	135	29.30	n.a.	100	Community	USA	White	MAT	.85
Lickenbrock and Braungart-Rieker (2015), male	135	30.79	n.a.	0	Community	USA	White	MAT	.85
Lin et al. (2017), female	141	39.84	n.a.	100	Community	TWN	Asian	KMSS	.96
Lin et al. (2017), male	141	42.06	n.a.	0	Community	TWN	Asian	KMSS	.96
Logan and Cobb (2013)	268	23.60	2.74	82	Student	CAN	Other	RAS	.87
Meltzer et al. (2014), Study 3, female	72	23.54	n.a.	100	Community	USA	White	QMI	.89
Meltzer et al. (2014), Study 3, male	72	24.92	n.a.	0	Community	USA	White	QMI	.89
Meltzer et al. (2013), female	169	23.40	n.a.	100	Community	USA	n.a.	QMI	.92
Meltzer et al. (2013), male	169	25.60	n.a.	0	Community	USA	n.a.	QMI	.92
Menéndez et al. (2011), female	108	27.12	n.a.	100	Community	ESP	n.a.	Ad-hoc	.81
Menéndez et al. (2011), male	79	28.73	n.a.	0	Community	ESP	n.a.	Ad-hoc	.81
Miller et al. (2003), female	168	21.00	n.a.	100	Community	USA	White	MOQ	.91
Miller et al. (2003), male	168	24.00	n.a.	0	Community	USA	White	MOQ	.91
Moen (2012), female	306	22.59	n.a.	100	Community	USA	White	KMSS	.86
Moen (2012), male	306	24.50	n.a.	0	Community	USA	White	KMSS	.87
Mund et al. (2015), Study 1	186	26.82	6.31	66	Community	DEU	White	RAS	.88
Mund et al. (2015), Study 2, female	2,124	31.35	9.38	100	National	DEU	White	RAS	.75
Mund et al. (2015), Study 2, male	2,124	34.16	9.38	0	National	DEU	White	RAS	.75
Naud et al. (2013), female	299	28.00	n.a.	100	Community	CAN	n.a.	DAS	.86
Naud et al. (2013), male	299	30.00	n.a.	0	Community	CAN	n.a.	DAS	.81
Nguyen et al. (2017), female	414	26.30	4.20	100	Community	USA	Hispanic	Ad-hoc	.70
Nguyen et al. (2017), male	414	27.90	4.20	0	Community	USA	Hispanic	Ad-hoc	.70
Niessen et al. (2018)	133	42.75	n.a.	31	Community	DEU	White	QMI	.97

Ogolsky et al. (2016), female	193	23.26	2.19	100	Community	USA	Other	MOQ	.75
Ogolsky et al. (2016), male	183	24.80	2.19	0	Community	USA	Other	MOQ	.75
Orth et al. (2015)	2,509	47.60	n.a.	40	Community	DEU	White	SRS	.93
Orth et al. (2012)	1,448	49.83	n.a.	57	Community	USA	White	RSS	.88
Padilla et al. (2018), female	246	39.38	19.12	100	Community	USA	Hispanic	CRDQ	.94
Padilla et al. (2018), male	246	41.94	19.12	0	Community	USA	Hispanic	CRDQ	.94
Paleari et al. (2005), female	124	43.80	22.60	100	Community	ITA	White	QMI	.96
Paleari et al. (2005), male	119	46.20	22.60	0	Community	ITA	White	QMI	.95
Parfitt et al. (2014), female	75	33.04	6.08	100	Community	GBR	White	DAS	.92
Parfitt et al. (2014), male	66	34.08	6.33	0	Community	GBR	White	DAS	.92
Parise et al. (2017), female	139	29.20	5.50	100	Community	ITA	White	QMI	.89
Parise et al. (2017), male	139	31.00	5.50	0	Community	ITA	White	QMI	.90
Peltz et al. (2018), female	249	35.00	10.50	100	Community	USA	White	CSI	.94
Peltz et al. (2018), male	249	36.00	10.50	0	Community	USA	White	CSI	.94
Reizer et al. (2014), Study 3, female	44	29.17	8.16	100	Community	ISR	n.a.	MAT	.84
Reizer et al. (2014), Study 3, male	44	30.34	8.16	0	Community	ISR	n.a.	MAT	.86
Roberson et al. (2015)	779	19.12	n.a.	81	Community	USA	White	CSI	.90
Robins et al. (2002)	214	21.00	2.08	63	National	NZL	White	Ad-hoc	.92
Ruffieux et al. (2014), female	162	40.40	14.60	100	Community	CHE	White	PFB	.92
Ruffieux et al. (2014), male	162	42.60	14.60	0	Community	CHE	White	PFB	.94
Sadikaj et al. (2015), female	93	27.89	4.22	100	Community	CAN	White	DAS	.93
Sadikaj et al. (2015), male	93	30.28	4.22	0	Community	CAN	White	DAS	.88
Schober (2012), female	5,624	33.55	7.94	100	Community	GBR	White	Ad-hoc	.80
Schober (2012), male	5,624	35.98	7.94	0	Community	GBR	White	Ad-hoc	.80
Sotskova et al. (2015), female	98	29.98	n.a.	100	Community	CAN	White	DAS	.91
Sotskova et al. (2015), male	98	32.03	n.a.	0	Community	CAN	White	DAS	.94
South et al. (2020), target	730	60.00	30.8	39	Community	USA	White	DAS	.82

South et al. (2020), spouse	551	60.00	30.8	61	Community	USA	White	DAS	.84
Sullivan et al. (2017)	86	22.50	n.a.	64	Community	USA	Other	RAS	.85
Sun et al. (2017), female	164	40.53	16.52	100	Community	USA	Black	CRDQ	.90
Sun et al. (2017), male	164	43.11	16.52	0	Community	USA	Black	CRDQ	.90
Szepsenwol et al. (2015), female	62	24.53	0.26	100	Community	ISR	n.a.	RAS	.82
Szepsenwol et al. (2015), male	62	25.87	0.26	0	Community	ISR	n.a.	RAS	.82
Tombeau Cost et al. (2018)	222	31.77	7.32	100	Community	CAN	n.a.	QMI	.81
Tremblay and Pierce (2011)	160	30.00	4.00	0	Community	CAN	White	DAS	.77
van den Troost et al. (2005), Study 3, female	386	45.00	25.80	100	National	NLD	White	Ad-hoc	.80
van den Troost et al. (2005), Study 3, male	386	47.50	25.80	0	National	NLD	White	Ad-hoc	.80
van Scheppingen et al. (2018)	84,711	29.59	n.a.	100	National	NOR	White	Ad-hoc	.88
Volling et al. (2015), female	241	31.60	9.57	100	Community	USA	White	IRQ	.76
Volling et al. (2015), male	241	33.20	9.57	0	Community	USA	White	IRQ	.75
Weidmann, Schönbrodt, et al. (2017), female	237	48.40	23.50	100	Community	CHE	White	RAS	.91
Weidmann, Schönbrodt, et al. (2017), male	237	50.70	23.50	0	Community	CHE	White	RAS	.91
Whiteman et al. (2007), female	188	36.74	17.36	100	Community	USA	White	CRDQ	.87
Whiteman et al. (2007), male	188	38.92	17.36	0	Community	USA	White	CRDQ	.87
Whitton et al. (2014)	748	25.71	3.14	65	Community	USA	White	DAS	.85
Zee and Weiss (2019)	678	46.41	20.70	50	National	USA	n.a.	Ad-hoc	.75

Note. Mean age and mean relationship duration are given in years. The column "Female" shows the percentage of female participants. T1 = Time 1. Reliability indicates the reliability estimate of the relationship-satisfaction measure. "n.a." indicates that data were not available. National = nationally representative. Country follows the ISO–3166–1 alpha–3 codes: CAN = Canada; CHE = Switzerland; CHN = China; DEU = Germany; ESP = Spain; FIN = Finland; GBR = United Kingdom of Great Britain and Northern Ireland; ITA = Italy; ISR = Israel; KOR = South Korea; NLD = Netherlands; NOR = Norway; NZL = New Zealand; USA = United States of America; TUR = Turkey; TWN = Taiwan. Measures were as follows (including adaptations, subscales, and translations thereof): CRDQ = Couple Relationship Domains Questionnaire (Huston et al., 1986); CSI = Couples Satisfaction Index (Funk & Rogge, 2007); DAS = Dyadic Adjustment Scale (Spanier, 1976); ENRICH = ENRICH Marital Satisfaction (Fowers & Olson, 1993); GMRS = Global Measure of Relationship Satisfaction (Lawrance & Byers, 1998); IRQ = Intimate Relations Questionnaire (Braiker & Kelley, 1979); KMSS = Kansas Marital Satisfaction Scale (Schumm et al., 1986); KMSS-R = Kansas Marital Satisfaction Scale Revised (Chung, 2004); MAT = Marital Adjustment Test (Locke & Wallace, 1959); MOQ = Marital Opinion Questionnaire (Huston & Vangelisti 1991); MSQO = Marital Satisfaction Questionnaire for Older Persons (Haynes et al., 1992); MRI = Marital Relationship Inventory (Burgess et al., 1971); PFB = Partnership Questionnaire (Partnerschaftsfragebogen; Hahlweg, 1996); PRQC = Perceived Relationship Quality Components Inventory (Fletcher et al., 2000); QMI = Quality of Marriage Index (Norton, 1983); RAS = Relationship Assessment Scale (Hendrick, 1988); REQ = Relationship Evaluation Questionnaire (Busby et al., 2001); RSI = Relationship Satisfaction subscale from the Investment Model Scale (Rusbult et al., 1998); RSS = Relationship Satisfaction Scale (Gilford & Bengtson, 1979); SRS = Self-Report Relationship Satisfaction (Schmitt et al., 1997); ad-hoc = measure constructed for the study (without a name).

Sample Characteristics

The publication year of the studies ranged from 2002 to 2019 (*Mdn* = 2014). Nine percent were nationally representative samples, 89% were community samples, and 2% were college samples. Almost all of the samples came from Western countries (96%), including the United States (47%), Canada (12%), the Netherlands (7%), Germany (6%), New Zealand (5%), Israel (4%), the United Kingdom (4%), Italy (3%), Switzerland (3%), Norway (2%), Spain (1%), Finland (1%), and Turkey (1%). Only 4% were from Asian countries (i.e., China, South Korea, and Taiwan). Data from African and South American countries were not available. Of the samples, 78% were predominantly White (with "predominantly" defined as at least 70% of the sample), 5% predominantly Asian, 3% predominantly Black, 3% predominantly Hispanic/Latin America, and 11% belonged to another ethnicity or were ethnically diverse. The mean proportion of female participants was 54% (range = 0% to 100%, *SD* = 47%, *Mdn* = 66%).

Participants' mean age at Time 1 was 34.78 years (*SD* = 10.36), ranging from 19.20 years to 71.00 years, and their mean relationship duration at Time 1 was 11.06 years (*SD* = 9.28), ranging from 3 months to 46.20 years. Participants' mean year of birth was 1969 (*SD* = 13.37), ranging from 1936 to 1995. Mean year of Time 1 assessment was 2004 (*SD* = 7.76), ranging from 1980 to 2014. The mean proportion of participants who lived in the same household as their partner was 82% (range = 11% to 100%, *SD* = 28%, *Mdn* = 100%), the mean proportion of participants who were married was 79% (range = 0% to 100%, *SD* = 33%, *Mdn* = 100%), and the mean proportion of participants who had children was 62% (range = 0% to 100%, *SD* = 44%, *Mdn* = 96%). As regards the occurrence of relationship transitions, 16% of the samples married or had a baby between Time 1 and any of the following assessments, and another 16% of the samples married or had a baby shortly before Time 1 (i.e., on average, 6.72 months before Time 1).

Methodological Characteristics

The mean time lag between assessments was 2.30 years (*SD* = 3.09), ranging from 2 months (which was the minimal lag required to be included in this meta-analysis) to 19.83 years. In 83% of the samples, relationship satisfaction was assessed with an established measure and in 17% with an ad-hoc measure. Of the studies with established measures, 46% used a global satisfaction measure and 54% an adjustment measure. A detailed list of all measures and their categorization into global satisfaction measures and adjustment measures is reported in Table S2.

Preliminary Analyses

Testing for Influential Outliers

We used the "influence" command in the metafor package (Viechtbauer, 2010) to test for influential outliers. In metafor, this command is not available for the "rma.mv" function, so we used the "rma" function, which does not account for the multilevel data structure. The results indicated that no sample qualified as potential outlier. Consequently, we used the full data set in the remainder of the analyses.

Testing for Publication Bias

We did not expect publication bias to be a problem in this meta-analysis given that most of the studies included did not focus on rank-order stability of relationship satisfaction per se (i.e., most studies focused on other research questions but the statistics that were central for this meta-analysis were reported in the studies).

For assessing publication bias, we used three methods. First, we examined the funnel plot, which shows the relation between effect size and standard error and serves as a graphical device to detect publication bias (Light & Pillemer, 1984; Rothstein et al., 2005; Sterne & Egger, 2001; Sutton, 2009). The funnel plot had an asymmetric shape (Figure 2). Second, we used Egger's regression test, which statistically tests for asymmetry of the funnel plot (Egger et al., 1997). In metafor, this test is not available for the "rma.mv" function, so we used the "rma" function. The test was not significant, $z = 0.755$, $p = .450$. Third, we compared effect

sizes that were published in the studies (k = 163) with effect sizes that were not published in the studies (but obtained from the authors upon request; k = 239), using a multilevel mixed-effects meta-regression model. The results indicated that effect sizes did not significantly differ from each other, Q_{Model} = 0.011, df = 1, p = .916. Overall, the findings from all three methods suggest that there was no evidence for systematic publication bias. In particular, comparing published versus unpublished effect sizes—which might be the most direct test of publication bias—indicated no significant differences.

Figure 2

Funnel Plot Displaying the Relation Between the Effect Size (Fisher's Z-Transformed Disattenuated Correlation Coefficient) and the Standard Error of the Effect Size

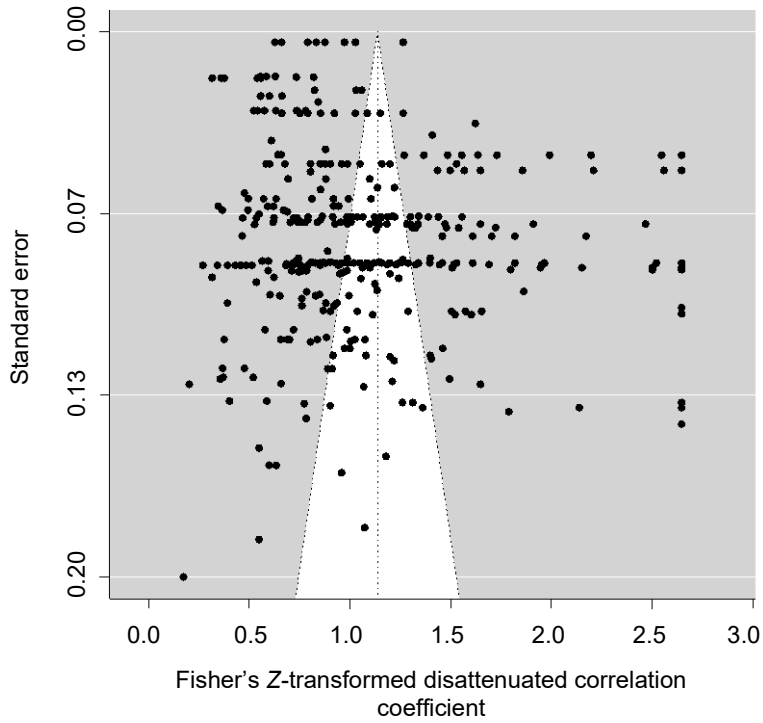
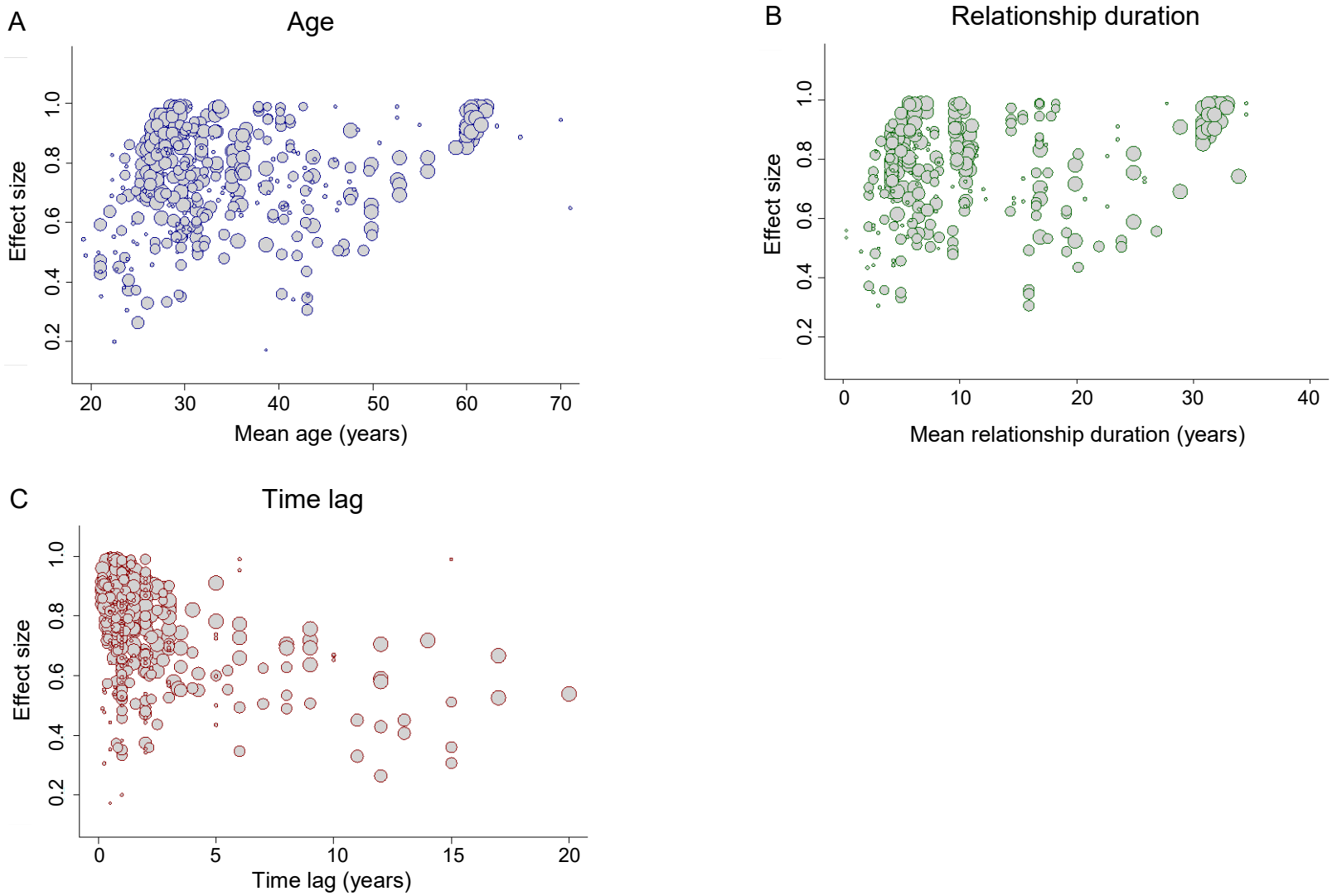


Figure 3

Scatterplots Displaying the Relation Between (A) Age and Effect Size, (B), Relationship Duration and Effect Size, and (C) Time Lag and Effect Size



Note. For effect size, the figure shows disattenuated correlation coefficients, indicating the rank-order stability of relationship satisfaction between two assessments. The size of the points reflects the weights of each sample.

Effect Size Analyses

Across samples, the weighted mean effect size was $r = .766$ with 95% CI [.732, .797]. When controlling for the effect of time lag, the weighted mean effect size was $r = .763$ with 95% CI [.728, .794]. Below, we report the effect size analyses for rank-order stability as a function of age and relationship duration.

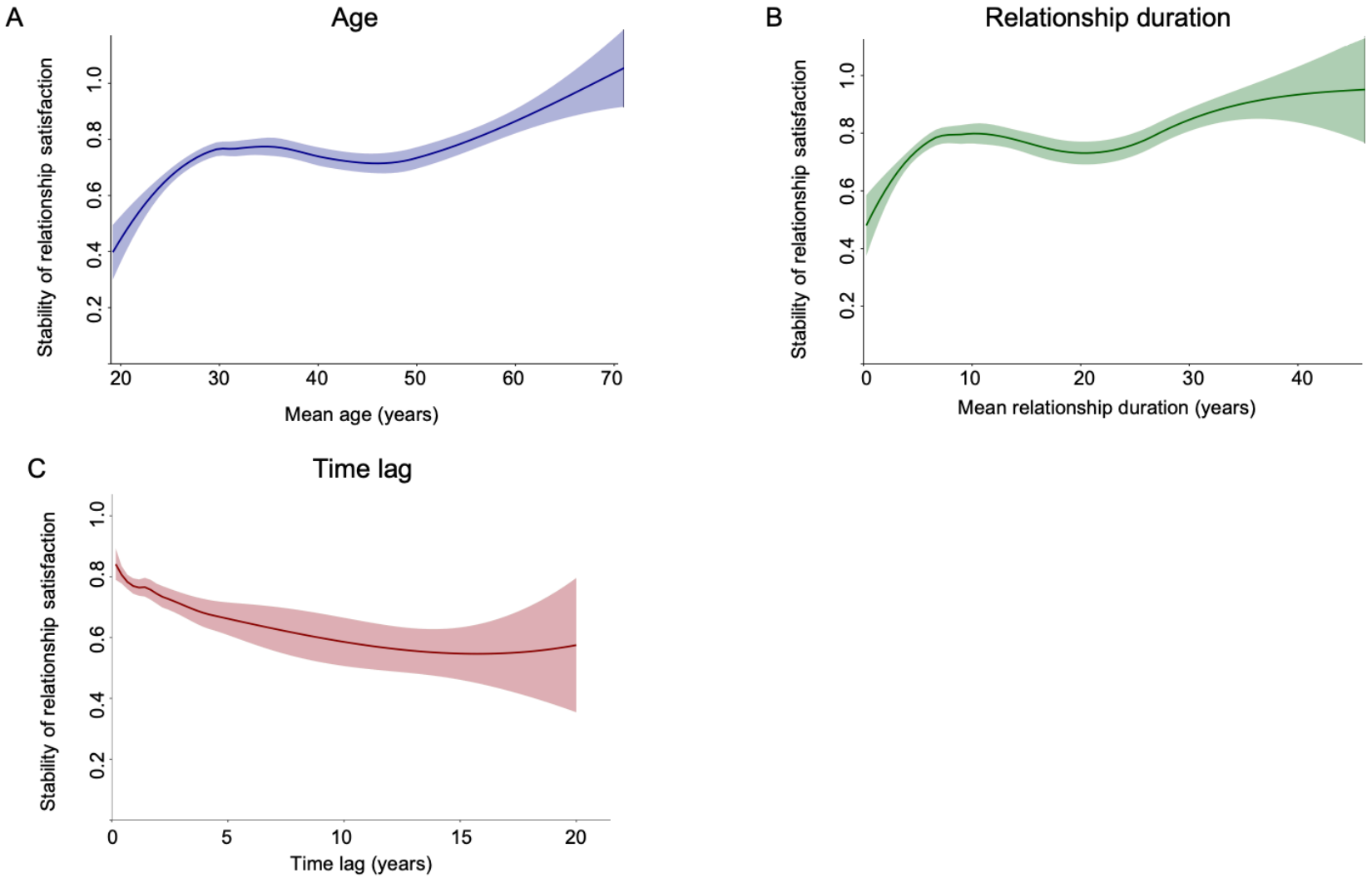
Rank-Order Stability as a Function of Age

Visual Overview. The scatterplot shown in Figure 3A provides an overview of the relation between mean age and effect size. As the figure indicates, variability of effect sizes was larger in young and middle adulthood compared to late adulthood. In addition, the scatterplot shows that the effect sizes became larger with age.

Moreover, we used the locally estimated scatterplot smoothing (LOESS) curve (Cleveland, 1978, 1981; Cleveland & Devlin, 1988) to gain further information about the developmental pattern of rank-order stability as a function of age. Figure 4A shows the LOESS curve with mean age as predictor and the disattenuated correlation coefficient as outcome. The figure supports the increasing trend of rank-order stability as a function of age, with a slight decline in rank-order stability between age 40 and 50 years.

Figure 4

LOESS Curves Showing the Rank-Order Stability of Relationship Satisfaction as a Function of (A) Age, (B) Relationship Duration, and (C) Time Lag



Note. The figure shows disattenuated correlation coefficients, indicating the rank-order stability of relationship satisfaction between two assessments, estimated with locally estimated scatterplot smoothing (LOESS) curves.

Weighted Mean Effect Sizes in Age Groups. For the meta-analytic computations, we constructed four age groups. For age 19 (19.20 years was the minimum age in the meta-analytic data set) to 50 years, we constructed age groups with 10-year intervals (except for the first group, which included 2 samples between 19 and 20 years): 19−30 years, 30−40 years, and 40−50 years. For age above 50 years, the number of samples was substantially lower. Therefore, we constructed one age group from 50−71 years (71 years was the maximum age in the meta-analytic data set).

Table S3 reports descriptive information on time lag in the four age groups of the effect size analyses (see values in the upper half of the table). As the table indicates, the mean time lag ranged from 1.88 to 4.24 years across age groups, with the largest mean time lag in the age group 40–50 years.

We used the weighted mean effect size as best estimate of rank-order stability of relationship satisfaction in each age group. Table 2 reports the meta-analytic estimates, both without control of time lag and with control of time lag. In the analyses that controlled for time lag, time lag was centered at the mean of the meta-analytic data set (i.e., 2.30 years). Thus, with control of time lag, differences between age groups in time lag are statistically controlled and, moreover, the meta-analytic estimates of rank-order stability refer to a time lag of 2.30 years. The findings from both analyses indicated that effect sizes were larger with higher age, except for the age group 40–50 years, which showed a slightly lower effect size compared to the age group 30–40 years. Overall, however, the findings suggest that rank-order stability of relationship satisfaction increased as a function of age, and this increase was particularly pronounced in young adulthood.

Table 2

Estimates of Rank-Order Stability of Relationship Satisfaction as a Function of Age and Relationship Duration

Group	*k*	*N*	Weighted mean effect size	95% CI	*Q*	Variances	
						σ_1^2	σ_2^2
Not controlled for time lag							
Age							
19–30 years	157	96,232	**.672**	[.613, .724]	**23737.5**	.065	.128
30–40 years	128	25,901	**.804**	[.751, .846]	**8827.2**	.185	.081
40–50 years	66	28,251	**.768**	[.696, .825]	**5139.7**	.186	.064
50–71 years	45	4,816	**.921**	[.865, .954]	**6290.4**	.134	.177
Relationship duration							
0–5 years	79	7,623	**.665**	[.597, .723]	**1472.7**	.070	.048
5–10 years	99	21,400	**.802**	[.733, .854]	**7326.9**	.151	.163
10–20 years	79	18,507	**.811**	[.736, .866]	**6884.1**	.249	.073
20–46 years	55	8,071	**.856**	[.757, .917]	**16828.9**	.296	.176
Controlled for time lag							
Age							
20–30 years	157	96,232	**.656**	[.596, .708]	**9659.4**	.061	.115
30–40 years	128	25,901	**.795**	[.741, .838]	**7937.4**	.176	.075
40–50 years	66	28,251	**.786**	[.715, .841]	**3333.4**	.195	.054
50–71 years	45	4,816	**.926**	[.863, .961]	**3005.2**	.226	.138
Relationship duration							
0–5 years	79	7,623	**.556**	[.457, .641]	**1375.9**	.087	.030
5–10 years	99	21,400	**.765**	[.686, .827]	**6620.3**	.140	.144
10–20 years	79	18,507	**.824**	[.761, .871]	**4286.3**	.195	.066
20–46 years	55	8,071	**.884**	[.785, .939]	**8470.8**	.370	.156

Note. Computations without control of time lag were made with multilevel random-effects models and computations with control of time lag were made with multilevel mixed-effects meta-regression models. In the analyses with time lag as covariate, time lag was grand-mean centered at 2.30 years. k = number of effect sizes; N = number of participants in each group. Weighted mean effect size = disattenuated correlation coefficient, indicating the rank-order stability of relationship satisfaction between two assessments. CI = confidence interval; Q = test statistic of the test for (residual) heterogeneity; σ_1^2 = variance proportion attributable to the level of the grouping variable (i.e., between samples); σ_2^2 = variance proportion attributable to the level nested within the grouping variable (i.e., within samples). Values in bold are significant at $p < .05$.

Rank-Order Stability as a Function of Relationship Duration

Visual Overview. The scatterplot shown in Figure 3B provides an overview of the relation between mean relationship duration and effect size. The variability of effect sizes was larger in relationships with a duration of less than 20 years compared to relationships with a duration of more than 20 years. The scatterplot also indicates that the effect sizes were larger with longer relationship duration.

Again, we generated a LOESS curve to gain further information about the developmental pattern of rank-order stability of relationship satisfaction as a function of relationship duration. Figure 4B shows the LOESS curve with mean relationship duration as predictor and the disattenuated correlation coefficient as outcome. The figure indicates an increasing trend of rank-order stability as a function of relationship duration, with a marginal decline after 20 years of relationship duration.

Weighted Mean Effect Sizes in Relationship-Duration Groups. For the meta-analytic computations, we constructed four relationship-duration groups. For 0 to 10 years of relationship duration, we constructed two groups with an interval of 5 years: 0−5 years and 5−10 years. For a relationship duration of more than 10 years, the number of samples was lower. Therefore, we constructed one group from 10 to 20 years and one group from 20 to 46 years (46 years was the maximum relationship duration in the meta-analytic data set).

Table S3 reports descriptive information on time lag in the four relationship-duration groups (see values in the lower half of the table). As the table indicates, the mean time lag ranged from 1.36 to 3.30 years across relationship-duration groups, with the largest mean in the group from 10 to 20 years.

Table 2 reports the meta-analytic estimates (both without control of time lag and with control of time lag). Again, in the analyses that controlled for time lag, time lag was centered at the mean of the meta-analytic data set (i.e., 2.30 years). In both sets of analyses, weighted mean effect sizes were larger with longer relationship duration, suggesting that rank-order

stability of relationship satisfaction increases as a function of relationship duration. This increase was particularly pronounced within the first ten years of the relationship.

Analyses With Time Lag

The scatterplot shown in Figure 3C provides an overview of the relation between time lag and effect size. Furthermore, Figure S1 shows the frequency distribution of time lag across samples. Although time lag ranged from 2 months to almost 20 years, most samples had a time lag of less than 5 years. The correlation between time lag and effect size was −.36 ($p < .001$), suggesting that rank-order stability was significantly lower over longer periods. To further examine rank-order stability of relationship satisfaction depending on time lag, we generated a LOESS curve. Figure 4C shows the LOESS curve with time lag as predictor and the disattenuated correlation coefficient as outcome. The figure indicates a decreasing trend of rank-order stability depending on time lag, suggesting that the curve levels off at an estimate of about .55.

Meta-Regressions Testing Age and Relationship Duration Simultaneously

Analyses based on age and relationship-duration groups are a useful method for describing the patterns of effects, but they might obscure the effects of continuous age and relationship duration. Moreover, age and relationship duration were highly correlated with each other (i.e., $r = .97$; see Table S4). Therefore, we tested multilevel mixed-effects meta-regression models using both time metrics as continuous predictors, to disentangle the effects of the two time metrics.

Specifically, we computed six models that differed systematically with regard to the predictors included. Model 1 tested the linear effect of age. Model 2 tested the linear effect of relationship duration. Model 3 tested the linear effects of age and relationship duration simultaneously. Model 4 tested the linear and quadratic effects of age. Model 5 tested the linear and quadratic effects of relationship duration. Finally, Model 6 tested the linear and

quadratic effects of age and relationship duration simultaneously. The findings from the meta-regression models are reported in Table 3, with control of time lag.

The findings showed that both linear and quadratic age (Models 1 and 4) and linear and quadratic relationship duration (Models 2 and 5) significantly predicted the effect sizes, consistent with the conclusions from the analyses with age and relationship-duration groups reported above. When linear age and linear relationship duration were tested simultaneously (Model 3), age was no longer significant, but relationship duration remained significant. Moreover, when quadratic age and quadratic relationship duration were included as additional predictors (Model 6), linear relationship duration remained a significant predictor of the effect sizes.

Overall, the meta-regression models allowed testing, with a different approach, rank-order stability as a function of age and relationship duration. The findings indicate that age and relationship duration significantly predicted rank-order stability of relationship satisfaction, suggesting greater stability with higher age and with longer relationship duration. Moreover, the results from the meta-regression models helped to disentangle the effect of age and relationship duration, suggesting that relationship duration was the more dominant (and significant) time metric for explaining rank-order stability of relationship satisfaction.

Table 3

Meta-Regression Predicting Effect Size from Age and Relationship Duration, Controlling for Time Lag

Predictor	Model					
	1	2	3	4	5	6
Linear age	**.257** [.168, .347]	—	−.140 [−.463, .183]	**.342** [.227, .457]	—	−.208 [−.544, .128]
Linear relationship duration	—	**.296** [.203, .390]	**.440** [.096, .785]	—	**.490** [.337, .643]	**.710** [.301, 1.119]
Quadratic age	—	—	—	**−.070** [−.130, −.010]	—	−.011 [−.142, .119]
Quadratic relationship duration	—	—	—	—	**−.114** [−.186, −.043]	−.109 [−.265, .046]

Note. $k = 306$. Computations were made with multilevel mixed-effects meta-regression models. Effect size = Fisher's Z-transformed disattenuated correlation coefficient, indicating the rank-order stability of relationship satisfaction between two assessments. The table shows unstandardized regression coefficients, with 95% confidence intervals in brackets. To avoid numerically small estimates, age and relationship duration were rescaled by the factor 10^{-1}. Moreover, in the present analyses age was centered at 35 years, and relationship duration was centered at 10 years. Dash indicates that this predictor is not included in the model. The analyses controlled for time lag, which was grand-mean centered at 2.30 years. Values in bold are significant at $p < .05$.

Moderator Analyses

Finally, we tested for moderators of the effect sizes, using multilevel mixed-effects meta-regression models. The zero-order correlations between the moderators, age, relationship duration, time lag, and effect size are provided in Table S4.

For the categorical moderator variables, we tested specific contrasts. As regards the occurrence of transitions, we contrasted samples that married or had a baby between Time 1 and any of the following occasions (16%) with samples that had no transition between Time 1 and any of the following occasions (84%). We also contrasted samples that married or had a baby shortly before Time 1 (16%) with samples that had no transition shortly before Time 1 (84%). As regards the measure of relationship satisfaction, we contrasted samples that were assessed with an ad-hoc measure (17%) versus established scale (83%). Among those samples that were assessed with an established scale, we contrasted global satisfaction scales (46%) with adjustment scales (54%). For sample type, we contrasted nationally representative samples (9%) with community and college samples (91%). For ethnicity, we contrasted samples that were predominantly White/European (78%) with samples that had another ethnicity (22%).

For the continuous variable baseline mean of relationship satisfaction, we had to take into account that the primary studies used different measures to assess relationship satisfaction, so that the observed Time 1 means were not directly comparable. Therefore, we converted the Time 1 means into POMP scores to make them comparable across studies (Cohen et al., 1999; see also Cerasoli, 2014). To compute POMP scores, we used the following formula given by Cohen et al. (1999)

$$\text{POMP} = \frac{\text{observed} - \text{minimum}}{\text{maximum} - \text{minimum}} \times 100,$$

where *observed* is the observed Time 1 mean of relationship satisfaction, *minimum* is the scale's minimum possible score, and maximum is the scale's maximum possible score. Thus, the POMP score reflects the Time 1 mean as percentage of the scale's maximum possible score, ranging from 0 to 100.

In the moderator analyses, we controlled for time lag and between-sample differences in mean age and mean relationship duration. Controlling for time lag, mean age, and mean relationship duration was essential given that effect sizes varied as a function of these variables, as reported above (see also Table S4). Given the substantial number of tests in the moderator analyses (i.e., 11 tests), we used the Bonferroni method and adjusted the significance level to $p < .0045$ (i.e., dividing .05 by the number of tests).

The results of the moderator analyses are shown in Table 4 and indicated that the occurrence of relationship transitions before Time 1 and sample type had significant moderator effects. Specifically, samples who experienced (vs. did not experience) a relationship transition shortly before Time 1 had a lower rank-order stability of relationship satisfaction. Moreover, nationally representative samples had a lower rank-order stability of relationship satisfaction than community or college samples. When we tested these two moderators simultaneously, the coefficients were very similar to the coefficients from the first step of the analyses and significant on the Bonferroni adjusted significance level.[5]

None of the other study characteristics were significant, which strengthens confidence in the robustness of the findings. That is, rank-order stability of relationship satisfaction did not significantly differ by household shared with partner, marital status, presence of children, the occurrence of relationship transitions between assessments, type of measure, ethnicity, gender, and baseline mean of relationship satisfaction.

[5] When testing the two moderators simultaneously, the coefficients were $B = -.280$, $SE = .090$, $p = .002$ (occurrence of relationship transitions shortly before Time 1) and $B = -.490$, $SE = .154$, $p = .002$ (sample type).

Table 4

Meta-Regression Models for Study Characteristics Predicting the Effect Size, Controlling for Mean Age, Mean Relationship Duration, and Time Lag

Moderator	*k*	*B*	*SE*	*p*
Household shared with partner	141	.285	.244	.243
Married	275	.298	.169	.078
Presence of children	207	.196	.162	.228
Transition	306	.087	.115	.451
Post transition	306	**−.276**	.091	.002
Measure (established)	306	.065	.125	.606
Measure (global)	249	−.209	.105	.047
Sample type	306	**−.476**	.153	.002
Ethnicity	291	.095	.120	.427
Female	306	−.030	.107	.779
Baseline mean	265	−.021	.009	.027

Note. Computations were made with multilevel mixed-effects meta-regression models. The effect size used was the Fisher's *Z*-transformed disattenuated correlation coefficient, indicating the rank-order stability of relationship satisfaction between two assessments. Mean age, mean relationship duration, and time lag were included as control variable in all models. Mean age, mean relationship duration, time lag, and mean year of birth were grand-mean centered prior to the analyses. Household shared with partner, married, presence of children, and female were proportions. The following variables were dichotomous: transition (1 = relationship transition between Time 1 and subsequent assessment, 0 = no relationship transition between Time 1 and subsequent assessment), post transition (1 = relationship transition shortly before Time 1 , 0 = no relationship transition shortly before Time 1), the two contrasts for measure of relationship satisfaction (1 = established scale, 0 = ad-hoc measure; 1 = global satisfaction, 0 = adjustment measure), sample type (1 = nationally representative sample, 0 = community or student sample), ethnicity (1 = predominantly White, 0 = other). *k* = number of effect sizes.

Discussion

The goal of this meta-analysis was to gain a robust and precise picture of rank-order stability of relationship satisfaction across adulthood, as a function of age and relationship duration. The meta-analytic data set was based on 148 independent samples, including 402 effect sizes from 153,396 participants. The mean age associated with the effect sizes ranged from 19 to 71 years, and the mean relationship duration ranged from 3 months to 46 years. The findings indicated that individual differences in relationship satisfaction within a given relationship are relatively stable over time (average r = .76, corrected for attenuation due to measurement error and based on an average time lag of 2.30 years). The findings also suggested that rank-order stability of relationship satisfaction varied systematically as a function of age and as a function of relationship duration. Specifically, rank-order stability of relationship satisfaction increased from young to late adulthood, with a slight decrease in middle adulthood. Moreover, rank-order stability of relationship satisfaction increased over the course of the relationship, with a slight decline around 20 years of relationship duration. In the analyses that examined age and relationship duration simultaneously, the findings suggested that relationship duration was the more dominant time metric for explaining rank-order stability of relationship satisfaction. Finally, the moderator analyses suggested that the occurrence of relationship transitions shortly before Time 1 and sample type explained variance in rank-order stability of relationship satisfaction. Except for these two moderators, the pattern of findings was robust across study characteristics such as marital status, ethnicity, and gender.

Rank-Order Stability of Relationship Satisfaction Across Adulthood

As noted in the Introduction, examining rank-order stability of relationship satisfaction concerns the question to which degree relationship satisfaction should be conceptualized as a trait-like construct (Fraley & Roberts, 2005). When comparing the present meta-analytic estimate of rank-order stability of relationship satisfaction within a given relationship (r =

.76) with findings on other psychological constructs, it is important to note that this estimate is disattenuated (i.e., corrected for unreliability of the measures used in individual studies) and, therefore, comparison estimates should also be corrected for unreliability. Moreover, comparison estimates should be based on a similar time lag (i.e., in the present research the average time lag was 2.30 years). The data from Anusic and Schimmack (2016) suggest that estimates of rank-order stability (corrected for measurement error and based on an average time lag of 2.30) are about .88 for personality traits, .80 for self-esteem, .75 for life satisfaction, and .88 for affect.

Thus, the present findings suggest that individual differences in relationship satisfaction are less stable than individual differences in personality traits, self-esteem, and affect, but as stable as individual differences in life satisfaction. Hence, similar to life satisfaction, relationship satisfaction can be considered, to some degree, as a trait-like construct. The relatively high rank-order stability of relationship satisfaction is particularly interesting given that relationship satisfaction—in contrast to life satisfaction—is not a pure characteristic of the individual, but is conceptually linked to the specific relationship environment, including characteristics of the romantic partner and of the relationship with the partner. It is likely that both individual predispositions (such as individual differences in emotional stability, self-esteem, and secure attachment; e.g., McNulty, 2016) and relationship characteristics (such as constructive communication patterns; e.g., Karney & Bradbury, 2020) contribute to the relatively high rank-order stability of relationship satisfaction. Clearly, future research using statistical models such as the STARTS model (Kenny & Zautra, 2001) is needed to more directly estimate stable and unstable variance components in relationship satisfaction (for an example of using the STARTS model, see Lucas & Donnellan, 2007). Studies using the STARTS model could also contribute to examining the asymptote of rank-order stability over long periods (Anusic & Schimmack, 2016; Fraley & Roberts, 2005; Kuster & Orth, 2013; Schuerger et al., 1989).

In this research, we found that the pattern of how rank-order stability of relationship satisfaction changed as a function of time lag was similar to that of other psychological constructs. More precisely, we observed that the rank-order stability of relationship satisfaction was significantly lower over longer periods (see Figure 4C). This decline could be expected based on research on other psychological constructs, which generally show a decrease in rank-order stability as the time lag increases (e.g. Fraley & Roberts, 2005). At the same time, we found that the LOESS curve did not approach zero over long periods (e.g., 15–20 years), but leveled off at an estimate of about .55. This finding suggests that there is an enduring component of individual differences in relationship satisfaction, again consistent with general findings on the stability of individual differences in other psychological constructs (e.g. Fraley & Roberts, 2005). To return to the example from the Introduction: The findings imply that Heather, Tom, and Mary will likely maintain their relative standing on relationship satisfaction at later occasions: Heather, who reports high relationship satisfaction (compared to Tom and Mary) at the first assessment, is likely to report high relationship satisfaction (compared to Tom and Mary) one year, five years, and even 15 years later. As discussed in the Introduction, it is possible that the long-term stability of relationship satisfaction can be explained by the long-term stability of the factors that influence couple members' relationship satisfaction. More specifically, the enduring component of individual differences in relationship satisfaction could be explained by characteristics of the individuals (e.g., personality characteristics such as the person's attachment orientation), the relationship (e.g., coping style of the couple), and the context (e.g., employment situation of the partners). Thus, although rank-order stability of relationship satisfaction levels off at a large value when assessed across long intervals, consistent with the notion of a relatively strong trait component of the construct, rank-order stability is likely due to influences from the person *and* influences from the relationship and context. As noted above, the present meta-analysis is based on data from assessments within a given relationship. Therefore, future research should test whether

the longterm estimate of rank-order stability holds when relationship satisfaction is assessed across different relationships of the same person.

At the same time, the present findings also showed that the degree of rank-order stability was still substantially below unity, indicating that there is room for change (Roberts & Nickel, 2021). To determine in which developmental periods change is more or less likely to occur, we examined rank-order stability of relationship satisfaction as a function of age and as a function of relationship duration.

Rank-Order Stability of Relationship Satisfaction as a Function of Age

The findings from both the LOESS curves and the group analyses indicated that, overall, rank-order stability of relationship satisfaction increased as a function of age, with a slight decline in middle adulthood. Specifically, the meta-analytic estimates from the group analyses were .66 for 19–30 years and .93 for 50–71 years. The estimates were not substantially altered when time lag between assessments was controlled for. The finding of greater rank-order stability with higher age is in line with the theoretical perspectives reviewed in the Introduction, according to which developmental tasks (Erikson, 1968; Havighurst, 1972) and selective investment (Carstensen et al., 1999; Charles & Carstensen, 2010) contribute to increasing rank-order stability of relationship satisfaction across adulthood.

The lower rank-order stability in young adulthood is consistent with research indicating that young adults, and in particular emerging adults, explore different life-path options, including romantic relationships (Arnett, 2000; Shulman & Connolly, 2013). Such exploration may contribute to less stable relationship conditions (e.g., young adults often are unmarried and do not live in a shared household; see also Table S4), which, in turn, could contribute to lower rank-order stability of relationship satisfaction in this life stage. However, in the present meta-analysis, being married and sharing a household did not moderate the effect sizes (see Table 4). Thus, even if young adult couples are more often unmarried and

live more often in separate households, this likely does not explain the lower rank-order stability in this developmental period.

Although the slight decline of rank-order stability in middle adulthood deviates from the general upward trend across adulthood, this tentative drop may be related to specific developmental challenges in midlife, which could affect romantic relationships in different ways. For example, many middle adults (but not all) spend more time with work than with leisure activities, and take on more generative, caring, and sometimes stressful social roles (e.g., Hudson et al., 2019). These greater challenges might deplete their resources to invest in the romantic relationship (Buck & Neff, 2012; Finkel & Campbell, 2001; Finkel et al., 2012; Finkel et al., 2014), leading to a temporary decline in the rank-order stability of relationship satisfaction in this developmental period. Nevertheless, we emphasize that the observed decline in rank-order stability in middle adulthood was relatively small.

Finally, the relatively large stability of relationship satisfaction in late adulthood is consistent with research suggesting that older adults invest more time and energy in positive social relationships with close others, such as with their relationship partner (Carstensen et al., 1999; Fredrickson & Carstensen, 1990; Fung et al., 1999). More investments, in turn, may lead to more stable relationship conditions, contributing to greater rank-order stability of relationship satisfaction in this life stage (Fraley & Roberts, 2005).

Rank-Order Stability of Relationship Satisfaction as a Function of Relationship Duration

The findings from both the LOESS curves and the group analyses indicated that rank-order stability of relationship satisfaction generally increased as a function of relationship duration, with a slight decline around 20 years of relationship duration. Specifically, the meta-analytic estimates from the group analyses were .56 in the first 5 years of relationships and .88 after a relationship duration of 20 years and longer. Again, the estimates were not substantially altered when time lag was controlled for. Interestingly, the confidence intervals were larger for couple members with long relationship duration (i.e., 35 years and more),

which may be due either to the lower number of samples who have been together for long periods or to greater heterogeneity of rank-order stability in samples with long relationship duration. Nevertheless, the overall trend suggested that rank-order stability generally increases over the course of relationships. This finding is in line with the theoretical perspectives reviewed in the Introduction, which suggested that in relationships that "survive" the first years, relationship conditions tend to be more stable, which consequently leads to increasing rank-order stability of relationship satisfaction as a function of relationship duration (Diekmann & Mitter, 1984; Kulu, 2014; Kurdek, 1998, 1999).

There are at least two mechanisms that could explain this phenomenon. First, the degree of rank-order stability depends on the stability of the context (see Fraley & Roberts, 2005). If a population as a whole (e.g., couples who have been together for 10 years or longer) is embedded in relatively consistent environments, this contributes to greater rank-order stability (Moss & Susman, 1980; Roberts & DelVecchio, 2000). By definition, in relationships of long duration, romantic partners have spent a considerable amount of time together and their relationship environment is likely characterized by accumulated consistency and stability. Of course, this does not mean that individuals in relationships of long duration do not experience any changes in their relationship environment, for example due to diseases or transitioning into retirement (e.g., Specht et al., 2011). Yet, it seems that these experiences do not generate substantial changes in the rank ordering of relationship satisfaction over time. Moreover, couples who have spent a considerable amount of time together often have invested many resources in their relationship, such as financial or emotional resources. According to the investment model of relationships (Rusbult et al., 1998), these investments (together with a relatively high level of relationship satisfaction and low quality of alternatives) maintain relationship commitment and may contribute to greater stability of the relationship. In addition, couples who are more satisfied in the relationship anticipate that they will remain in the relationship and hence invest in more adaptive

relationship processes (e.g., constructive coping styles), strengthening their high level of satisfaction. Similarly, less satisfied couples may not believe in continuing the relationship and will invest much less in relationship-promoting behavior, which maintains their relatively low level of satisfaction. Thus, these relationship processes may explain why individual differences in relationship satisfaction are relatively stable over time.

Second, identity processes could also account for greater rank-order stability of relationship satisfaction, combining psychological and environmental mechanisms (Roberts & DelVecchio, 2000). Only recently, the concept of narrative identity has been applied to the romantic relationship domain, illustrating that romantic relationships—similar to whole lives (McAdams, 1995, 2013)—can be represented by means of narrative identity (Bühler & Dunlop, 2019). Specifically, the story that individuals form about their relationship reflects their relationship identity. As argued in the context of individual lives, people with a strong identity make decisions and choose life paths that are consistent with their personality and their identity (see selective person-environment transactions; Caspi, 1998). Having a strong sense of identity is also linked to continued investment in the chosen life path, which predicts an increase in ego-resilience, that is, a better ability for adjusting to changes in the environment and recovery from difficult situations (Pals, 1999). Moreover, a strong sense of identity also exerts a filter function as to which information is perceived and processed by the person. Finally, once a person's identity has become known to others, they create reputations about the person (Hogan & Roberts, 2000) and respond in a way that strengthens the person's personality and identity (see evocative person-environment transactions; Caspi, 1998). Together, these identity processes may contribute to greater stability of individual differences in psychological constructs (Roberts & DelVecchio, 2000).

Applied to the context of romantic relationships, this means that individuals with a strong relationship identity might make decisions and choose relationship paths that are in line with their relationship identity (e.g., seeing a couple therapist when relationship problems

arise). Having a strong relationship identity may also be linked to continuous investment in the relationship, which contributes to a resilient and stable relationship. Moreover, people with a strong relationship identity may selectively perceive and process information in a way that strengthens their relationship identity (e.g., focusing on positive, rather than negative, attributes of the relationship or the relationship partner). Finally, the reputation of the relationship (e.g., "Max and Alex have been happily married ever since") may feedback into the relationship so that others react to the couple in a way that is consistent with their established relationship identity. Clearly, the same relationship-identity mechanisms apply to less satisfied couples. For instance, individuals who are less satisfied with their relationship may make decisions that are in line with their relationship identity (e.g., "We do not fit together anyway") and may not consider, or even decline, to see a couple therapist when problems arise.

At the same time, meta-analytic findings on mean levels suggest that relationship satisfaction is, on average, relatively high over the course of the relationship, never falling below 77% of the maximum possible (Bühler et al., 2021). These findings also correspond with research based on latent class growth analysis and group-based modeling, indicating that the majority of couples (around 70–90%) remains fairly satisfied over time, while only a subgroup of couples (around 10–30%) experiences greater declines in relationship satisfaction and potentially a relationship breakup (e.g., Anderson et al., 2010; Birditt et al., 2012; Foran et al., 2013; Lavner & Bradbury, 2010; Lavner et al., 2012; Lorber et al., 2015). Thus, among intact couples, even the less satisfied couple members are often relatively satisfied, which may explain why they remain in the relationship over longer periods of time. In addition, as described in the investment model, the level of relationship satisfaction is only one of the factors that contributes to commitment and stability in the relationship. Hence, among the intact, less satisfied couples there may be other factors, such as the investments that couple members have already made, that motivate partners to stay together. Together, all of these

mechanisms may stabilize individual differences in relationship satisfaction and contribute to greater rank-order stability of relationship satisfaction over the course of the relationship.

The Role of Age and Relationship Duration

As noted in the Introduction, age and relationship duration are generally highly correlated, which was also the case in the present meta-analysis (r = .97; see Table S4). This high correlation, however, is not a methodological artifact, but reflects the reality of most couples: The older people are, the longer they have typically been in their relationship, simply because they are older. Nevertheless, it is important to note that although the two time metrics are highly correlated, they are not interchangeable (similar to height and weight, which are strongly correlated but by no means the same construct; Fincham et al., 2018). Therefore, we tested multilevel mixed-effects meta-regression models using both time metrics as continuous predictors, with the goal of disentangling the effects of the two time metrics. The findings suggested that relationship duration, rather than age, was the more dominant time metric for explaining rank-order stability of relationship satisfaction within a given relationship. As reviewed in the Introduction, Fraley and Roberts' (2005) model suggests that stable environments contribute to higher rank-order stability of individual-difference constructs. The more dominant role of relationship duration for rank-order stability of relationship satisfaction might be explained by the more stable relationship environments that longterm couples usually experience. This, in turn, may result in more stable transactions between couple members and their relationship environments (Fraley & Roberts, 2005). With longer relationship duration, relationship partners also tend to become more congruent in personality traits such as agreeableness, conscientiousness, and openness (Rammstedt & Schupp, 2008), which may further stabilize the transactions between the partners and their environment over the course of their relationship.

At the same time, we emphasize that it was not possible to unequivocally isolate the unique effects of the time metrics because the time metrics are, as described above, inherently

linked to each other. Thus, in studies with romantic couples both age and relationship duration should be examined simultaneously to account for both of their effects. Otherwise, developmental processes in romantic relationships might mistakenly be attributed to one of the time metrics, simply because the other time metric was not examined in the research.

Moderators of the Rank-Order Stability of Relationship Satisfaction

The pattern of findings was relatively robust across the moderators tested, with two exceptions. First, in line with our hypotheses the occurrence of relationship transitions shortly before Time 1 predicted lower rank-order stability of relationship satisfaction. Relationship transitions, such as the birth of a child or marriage, might confront people with new demands and challenges and might destabilize relationship environments (Doss et al., 2009; Fraley & Roberts, 2005; Sanchez & Thomson, 1997). A destabilized relationship environment, in turn, may lead to lower rank-order stability of relationship satisfaction. Second, nationally representative samples had a lower rank-order stability of relationship satisfaction than community samples and samples of college students. As noted in the Introduction, we had no hypotheses about the moderating effect of sample type on the effect size. The present findings suggest that romantic relationships in community and college samples may be more stable than romantic relationships in nationally representative samples. However, more research is needed for drawing stronger conclusions about the moderating effects of relationship transitions and sample type on the rank-order stability of relationship satisfaction.

A general conclusion from the moderator analyses is that the findings were robust across most sample and methodological characteristics. In other words, rank-order stability of relationship satisfaction did not significantly differ in samples with different compositions with regard to shared household, marital status, presence of children, ethnicity, and gender. Moreover, the findings did not differ in samples that had versus had not undergone relationship transitions between Time 1 and any of the following measurement occasions, in samples with different measures and with different baseline means of relationship satisfaction.

Limitations and Future Directions

Although this meta-analysis included data from 153,396 participants from 16 countries, one limitation is that the majority of samples were from Western countries (e.g., United States, Canada, the Netherlands, Germany). Given that the meta-analytic data set included few studies from Asian countries and none from African and South American countries, the present data did not allow testing whether the results hold outside of North America and Europe. Moreover, nearly all samples consisted of participants from opposite-sex relationships, and only one sample included participants from same-sex relationships. Therefore, more primary studies are needed that examine couples from different cultural backgrounds (Henrich et al., 2010) and individuals involved in diverse types of relationships, including same-sex relationships (e.g., Chen & van Ours, 2018).

Furthermore, the meta-analytic data set might be selective to some degree because samples were only included if they provided data on at least two measurement occasions (thus, couples who separated after the first assessment did not provide the second assessment needed for computing rank-order stability). To reduce concerns related to this limitation, we used a relatively short time lag between assessments (i.e., 2 months). Nevertheless, these "surviving" relationships might still represent relationships that are above average in their commitment given that dissolving relationships often show higher levels of stress and dissimilarity, both at baseline and over time (Finn et al., 2020). Therefore, the present findings should be interpreted in light of these more stable relationship environments, and samples of to-be-dissolved couples might show lower degrees of rank-order stability of relationship satisfaction. Similarly, the present data allow conclusions about rank-order stability of relationship satisfaction within a given relationship but cannot speak to rank-order stability of relationship satisfaction across relationships. Research that follows individuals within and across relationships would generate knowledge about such different degrees of rank-order stability and, moreover, provide important knowledge on the trait-like nature of

relationship satisfaction across relationships (for cross-relationship research, see Johnson & Neyer, 2019).

Another limitation is that the number of samples was relatively low for samples with a mean age of above 50 years and with a relationship duration of above 20 years. Consequently, the conclusions for rank-order stability of relationship satisfaction are less strong for these developmental and relationship periods. In addition, there were no effect sizes available for samples older than 71 years, which could be explained, at least partially, by health constraints in this age group or by a higher likelihood that one of the partners was already deceased. Future research is needed that specifically studies romantic relationships and relationship trajectories in middle adulthood and especially late adulthood, including "gray divorces" (Brown & Lin, 2012; Gloor et al., 2021). Finally, studying newlyweds in late adulthood would contribute to lowering the correlation between age and relationship duration and would provide further insights into the rank-order stability of relationship satisfaction as a function of both age and relationship duration.

Conclusion

The present meta-analysis examined the rank-order stability of relationship satisfaction, synthesizing data from 148 samples with more than 150,000 participants. Overall, the findings showed that individual differences in relationship satisfaction are relatively stable within a given relationship. The moderator analyses strengthened confidence in this finding, given that rank-order stability was robust across most sample and methodological characteristics. Moreover, the findings indicated that rank-order stability of relationship satisfaction changes systematically as a function of age and of relationship duration. For both time metrics, there was an overall increase of rank-order stability over time. As regards age, rank-order stability increased from young to late adulthood, with a slight decrease in middle adulthood. As regards relationship duration, rank-order stability generally increased over the course of the relationship, with a slight decline around 20 years

of relationship length. When both time metrics were examined simultaneously, relationship duration tended to be the more dominant time metric for explaining rank-order stability of relationship satisfaction. Taken together, the present findings may stimulate future research on the developmental processes that underlie stability and change of relationship satisfaction across adulthood.

References

References marked with an asterisk indicate studies included in the meta-analysis.

Anderson, J. R., Van Ryzin, M. J., & Doherty, W. J. (2010). Developmental trajectories of marital happiness in continuously married individuals: A group-based modeling approach. *Journal of Family Psychology, 24*(5), 587–596. https://doi.org/10.1037/a0020928

*Andres, M. (2014). Distress, support, and relationship satisfaction during military-induced separations: A longitudinal study among spouses of Dutch deployed military personnel. *Psychological Services, 11*(1), 22–30. https://doi.org/10.1037/a003375022

Anusic, I., & Schimmack, U. (2016). Stability and change of personality traits, self-esteem, and well-being: Introducing the meta-analytic stability and change model of retest correlations. *Journal of Personality and Social Psychology, 110*(5), 766–781. https://doi.org/https://doi.org/10.1037/pspp0000066

Appelbaum, M., Cooper, H., Kline, R. B., Mayo-Wilson, E., Nezu, A. M., & & Rao, S. M. (2018). Journal article reporting standards for quantitative research in psychology: The APA Publications and Communications Board task force report. *American Psychologist, 73*(1), 3–25. https://doi.org/10.1037/amp0000191

Ardelt, M. (2000). Still stable after all these years? Personality stability theory revisited. *Social Psychology Quarterly, 63*(4), 392–405. https://doi.org/10.2307/2695848

Arnett, J. J. (2000). Emerging adulthood: A theory of development from the late teens through the twenties. *American Psychologist, 55*(5), 469–480. https://doi.org/10.1037/0003-066x.55.5.469

*Be, D., Whisman, M. A., & Uebelacker, L. A. (2013). Prospective associations between marital adjustment and life satisfaction. *Personal Relationships, 20*(4), 728–739. https://doi.org/10.1111/pere.12011

Belsky, J., & Rovine, M. (1990). Patterns of marital change across the transition to parenthood: Pregnancy to three years postpartum. *Journal of Marriage and the Family, 52*(1), 5–19. https://doi.org/10.2307/352833

*Bikos, L. H., Çiftçi, A., Güneri, O. Y., Demir, C. E., Sümer, Z. H., Danielson, S., DeVries, S., & Bilgen, W. A. (2007). A repeated measures investigation of the first-year adaptation experiences of the female expatriate spouse living in Turkey. *Journal of Career Development, 34*(1), 5–26. https://doi.org/10.1177/0894845307304063

Birditt, K. S., Hope, S., Brown, E., & Orbuch, T. (2012). Developmental trajectories of marital happiness over 16 years. *Research in Human Development, 9*(2), 126–144. https://doi.org/10.1080/15427609.2012.680844

*Bloch, L., Haase, C. M., & Levenson, R. W. (2014). Emotion regulation predicts marital satisfaction: More than a wives' tale. *Emotion, 14*(1), 130–144. https://doi.org/10.1037/a0034272

Block, J. (2014). *Lives through time*. Psychology Press.

*Blumenstock, S. M., & Papp, L. M. (2017). Sexual distress and marital quality of newlyweds: An investigation of sociodemographic moderators. *Family Relations: An Interdisciplinary Journal of Applied Family Studies, 66*(5), 794–808. https://doi.org/10.1111/fare.12285

Bodenmann, G. (1995). A systemic-transactional conceptualization of stress and coping in couples. *Swiss Journal of Psychology, 54*(1), 34–49.

*Bodi, O., Mikula, G., & Riederer, B. (2010). Long-term effects between perceived justice in the division of domestic work and women's relationship satisfaction: Testing for moderation effects. *Social Psychology, 41*(2), 57–65. https://doi.org/10.1027/1864-9335/a000009

*Bouchard, G. (2014). The quality of the parenting alliance during the transition to parenthood. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement, 46*(1), 20–28. https://doi.org/10.1037/a003125920

*Bouchard, G., Boudreau, J., & Hébert, R. (2006). Transition to parenthood and conjugal life: Comparisons between planned and unplanned pregnancies. *Journal of Family Issues, 27*(11), 1512–1531. https://doi.org/10.1177/0192513X06290855

*Bower, D., Jia, R., Schoppe-Sullivan, S. J., Mangelsdorf, S. C., & Brown, G. L. (2013). Trajectories of couple relationship satisfaction in families with infants: The roles of parent gender, personality, and depression in first-time and experienced parents. *Journal of Social and Personal Relationships, 30*(4), 389–409. https://doi.org/10.1177/0265407512457656

Bowlby, J. (1973). *Attachment and loss: Vol. II. Separation: Anxiety and anger*. Basic Books.

*Brown, B. J., Robinson, D., Jensen, J. F., Seedall, R. B., Hodgson, J., & Norton, M. C. (2019). Will improving my marriage improve my sleep? *Journal of Couple & Relationship Therapy, 18*(2), 85–103. https://doi.org/10.1080/15332691.2017.1417938

Brown, S. L., & Lin, I.-F. (2012). The gray divorce revolution: Rising divorce among middle-aged and older adults, 1990–2010. *The Journals of Gerontology: Series B, 67*(6), 731–741. https://doi.org/10.1093/geronb/gbs089

Buck, A. A., & Neff, L. A. (2012). Stress spillover in early marriage: The role of self-regulatory depletion. *Journal of Family Psychology, 26*(5), 698–708. https://doi.org/10.1037/a0029260

Bühler, J. L., & Dunlop, W. L. (2019). The narrative identity approach and romantic relationships. *Social and Personality Psychology Compass, 13*(4), Article e12447. https://doi.org/10.1111/spc3.12447

Bühler, J. L., Krauss, S., & Orth, U. (2021). Development of relationship satisfaction across the life span: A systematic review and meta-analysis. *Psychological Bulletin, 147*(10), 1012–1053. https://doi.org/http://dx.doi.org/10.1037/bul0000342

Bühler, J. L., Weidmann, R., Nikitin, J., & Grob, A. (2019). A closer look at life goals across adulthood: Applying a developmental perspective to content, dynamics, and outcomes of goal importance and goal attainability. *European Journal of Personality, 33*(3), 359–384. https://doi.org/10.1002/per.2194

*Busby, D. M., & Gardner, B. C. (2008). How do I analyze thee? Let me count the ways: Considering empathy in couple relationships using self and partner ratings. *Family Process, 47*(2), 229–242. https://doi.org/10.1111/j.1545-5300.2008.00250.x

*Buyukcan-Tetik, A., Campbell, L., Finkenauer, C., Karremans, J. C., & Kappen, G. (2017). Ideal standards, acceptance, and relationship satisfaction: Latitudes of differential effects. *Frontiers in Psychology, 8*, 1691. https://doi.org/10.3389/fpsyg.2017.01691

*Byers, E. S. (2005). Relationship satisfaction and sexual satisfaction: A longitudinal study of individuals in long-term relationships. *Journal of Sex Research, 42*(2), 113–118. https://doi.org/10.1080/00224490509552264

Campbell, L., Simpson, J. A., Boldry, J., & Kashy, D. A. (2005). Perceptions of Conflict and Support in Romantic Relationships: The Role of Attachment Anxiety. *Journal of Personality and Social Psychology, 88*(3), 510–531. https://doi.org/https://doi.org/10.1037/0022-3514.88.3.510

Carr, D., & Utz, R. L. (2020). Families in later life: A decade in review. *Journal of Marriage and Family, 82*(1), 346–363. https://doi.org/10.1111/jomf.12609

Carstensen, L. L., Isaacowitz, D. M., & Charles, S. T. (1999). Taking time seriously: A theory of socioemotional selectivity. *American Psychologist, 54*(3), 165–181. https://doi.org/10.1037/0003-066x.54.3.165

Caspi, A. (1998). Personality development across the life course. In W. Damon & N. Eisenberg (Eds.), *Handbook of child psychology: Social, emotional, and personality development* (pp. 311–388). John Wiley & Sons Inc.

Caspi, A., & Bem, D. J. (1990). Personality continuity and change across the life course. In L. A. Pervin (Ed.), *Handbook of personality: Theory and research* (pp. 549–575). Guilford Press.

Caspi, A., Bem, D. J., & Elder, G. H., Jr. (1989). Continuities and consequences of interactional styles across the life course. *Journal of Personality, 57*(2), 375–406. https://doi.org/10.1111/j.1467-6494.1989.tb00487.x

Caspi, A., & Roberts, B. W. (1999). Personality continuity and change across the life course. In *Handbook of personality: Theory and research* (pp. 300–326). Guilford Press.

Cast, A. D. (2004). Well-being and the transition to parenthood: An identity theory approach. *Sociological perspectives, 47*(1), 55–78. https://doi.org/10.1525/sop.2004.47.1.55

Cerasoli, C. P., Nicklin, J. M., & Ford, M. T. (2014). Intrinsic motivation and extrinsic incentives jointly predict performance: A 40-year meta-analysis. *Psychological Bulletin, 140*(4), 980–1008. https://doi.org/10.1037/a0035661

Charles, S. T., & Carstensen, L. L. (2010). Social and emotional aging. *Annual Review of Psychology, 61*(1), 383–409. https://doi.org/10.1146/annurev.psych.093008.100448

Chen, S., & van Ours, J. C. (2018). Subjective Well-being and Partnership Dynamics: Are Same-Sex Relationships Different? *Demography, 55*(6), 2299–2320. https://doi.org/10.1007/s13524-018-0725-0

*Choi, E. (2016). Marital satisfaction and maternal depressive symptoms among Korean mothers transitioning to parenthood. *Journal of Family Psychology, 30*(4), 516–521. https://doi.org/10.1037/fam0000178

*Christopher, C., Umemura, T., Mann, T., Jacobvitz, D., & Hazen, N. (2015). Marital quality over the transition to parenthood as a predictor of coparenting. *Journal of Child and Family Studies, 24*(12), 3636–3651. https://doi.org/10.1007/s10826-015-0172-0

Cleveland, W. S. (1978). Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association 74*(368), 829–836. https://doi.org/10.1080/01621459.1979.10481038

Cleveland, W. S. (1981). LOWESS: A Program for smoothing scatterplots by robust locally weighted regression. *The American Statistician, 35*(1), 54. https://doi.org/10.2307/2683591

Cleveland, W. S., & Devlin, S. J. (1988). Locally weighted regression: An approach to regression analysis by local fitting. *Journal of the American Statistical Association, 83*(403), 596−610. https://doi.org/10.1080/01621459.1988.10478639

Cohen, P., Cohen, J., Aiken, L. S., & West, S. G. (1999). The problem of units and the circumstance for POMP. *Multivariate Behavioral Research, 34*(3), 315–346. https://doi.org/10.1207/s15327906mbr3403_2

Conger, R. D., Conger, K. J., & Martin, M. J. (2010). Socioeconomic status, family processes, and individual development. *Journal of Marriage and Family, 72*(3), 685–704. https://doi.org/10.1111/j.1741-3737.2010.00725.x

*Crocker, J., Canevello, A., & Lewis, K. A. (2017). Romantic relationships in the ecosystem: Compassionate goals, nonzero-sum beliefs, and change in relationship quality. *Journal of Personality and Social Psychology, 112*(1), 58-75. https://doi.org/10.1037/pspi0000076

Debrot, A., Cook, W. L., Perrez, M., & Horn, A. B. (2012). Deeds matter: Daily enacted responsiveness and intimacy in couples' daily lives. *Journal of Family Psychology, 26*(4), 617–627. https://doi.org/https://doi.org/10.1037/a0028666

*DeMaris, A. (2010). The 20-year trajectory of marital quality in enduring marriages: Does equity matter? *Journal of Social and Personal Relationships, 27*(4), 449–471. https://doi.org/10.1177/0265407510363428

Diekmann, A., & Mitter, P. (1984). A comparison of the "sickle function" with alternative stochastic models of divorce rates. In A. Diekmann & P. Mitter (Eds.), *Stochastic modelling of social processes* (pp. 123–153). Academic Press.

*Doohan, E.-A. M., Carrere, S., & Riggs, M. L. (2010). Using relational stories to predict the trajectory toward marital dissolution: The oral history interview and spousal feelings of flooding, loneliness, and depression. *Journal of Family Communication, 10*(1), 57–77. https://doi.org/10.1080/15267430903396401

Doss, B. D., Rhoades, G. K., Stanley, S. M., & Markman, H. J. (2009). The effect of the transition to parenthood on relationship quality: An 8-year prospective study. *Journal of Personality and Social Psychology, 96*. https://doi.org/10.1037/a0013969

Ebner, N. C., Freund, A. M., & Baltes, P. B. (2006). Developmental changes in personal goal orientation from young to late adulthood: From striving for gains to maintenance and prevention of losses. *Psychology and Aging, 21*(4), 664–678. https://doi.org/10.1037/0882-7974.21.4.664.

Egger, M., Davey Smith, G., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal, 315*(7109), 629–634. https://doi.org/10.1136/bmj.315.7109.629

Erikson, E. H. (1968). *Identity: Youth and crisis*. Norton.

*Erol, R. Y., & Orth, U. (2014). Development of self-esteem and relationship satisfaction in couples: Two longitudinal studies. *Developmental Psychology, 50*(9), 2291–2303. https://doi.org/10.1037/a0037370

Erol, R. Y., & Orth, U. (2016). Self-esteem and the quality of romantic relationships. *European Psychologist, 21*(4), 274–283. https://doi.org/10.1027/1016-9040/a000259

*Fallis, E. E., Rehman, U. S., Woody, E. Z., & Purdon, C. (2016). The longitudinal association of relationship satisfaction and sexual satisfaction in long-term relationships. *Journal of Family Psychology, 30*(7), 822–831. https://doi.org/10.1037/fam0000205

*Fincham, F. D., & Beach, S. R. H. (2007). Forgiveness and marital quality: Precursor or consequence in well-established relationships? *The Journal of Positive Psychology, 2*(4), 260–286. https://doi.org/10.1080/17439760701552360

Fincham, F. D., Rogge, R., & Beach, S. R. H. (2018). Relationship satisfaction. In A. L. Vangelisti & D. Perlman (Eds.), *The Cambridge handbook of personal relationships* (pp. 422–436). Cambridge University Press. https://doi.org/10.1017/9781316417867.033

Finkel, E. J., & Campbell, W. K. (2001). Self-control and accommodation in close relationships: An interdependence analysis. *Journal of Personality and Social Psychology, 81*(2), 263–277. https://doi.org/10.1037/0022-3514.81.2.263

Finkel, E. J., DeWall, C. N., Slotter, E. B., McNulty, J. K., Pond, R. S., Jr., & Atkins, D. C. (2012). Using I3 theory to clarify when dispositional aggressiveness predicts intimate partner violence perpetration. *Journal of Personality and Social Psychology, 102*(3), 533–549. https://doi.org/10.1037/a0025651

Finkel, E. J., Hui, C. M., Carswell, K. L., & Larson, G. M. (2014). The suffocation of marriage: Climbing Mount Maslow without enough oxygen. *Psychological Inquiry, 25*(1), 1–41. https://doi.org/10.1080/1047840x.2014.863723

Finn, C., Johnson, M. D., & Neyer, F. J. (2020). Happily (n)ever after? Codevelopment of romantic partners in continuing and dissolving unions. *Developmental Psychology, 56*(5), 1022-1028. https://doi.org/10.1037/dev0000897

Finn, C., Mitte, K., & Neyer, F. J. (2013). The relationship-specific interpretation bias mediates the link between neuroticism and satisfaction in couples. *European Journal of Personality, 27*(2), 200–212. https://doi.org/10.1002/per.1862

Fisher, R. A. (1921). On the "probable error" of a coefficient of correlation deduced from a small sample. *Metron, 1*, 1–32.

Foran, H. M., Hahlweg, K., Kliem, S., & O'Leary, K. D. (2013). Longitudinal patterns of relationship adjustment among German parents. *Journal of Family Psychology, 27*(5), 838–843. https://doi.org/10.1037/a0034183

Fraley, R. C., & Roberts, B. W. (2005). Patterns of continuity: A dynamic model for conceptualizing the stability of individual differences in psychological constructs across the life course. *Psychological Review, 112*(1), 60–74. https://doi.org/10.1037/0033-295x.112.1.60

Frederick, D. A., Lever, J., Gillespie, B. J., & Garcia, J. R. (2017). What keeps passion alive? Sexual satisfaction is associated with sexual communication, mood setting, sexual variety, oral sex, orgasm, and sex frequency in a national U.S. Study. *Journal of Sex Research, 54*(2), 186–201. https://doi.org/10.1080/00224499.2015.1137854

Fredrickson, B. L., & Carstensen, L. L. (1990). Choosing social partners: How old age and anticipated endings make people more selective. *Psychology and Aging, 5*(3), 335–347. https://doi.org/10.1037/0882-7974.5.3.335

Freund, A. M. (2008). Successful aging as management of resources: The role of selection, optimization, and compensation. *Research in Human Development, 5*(2), 94–106. https://doi.org/10.1080/15427600802034827

Freund, A. M., & Nikitin, J. (2012). Junges und mittleres Erwachsenenalter. In W. Schneider & U. Lindenberger (Eds.), *Entwicklungspsychologie* (pp. 259–282). Beltz.

Freund, A. M., & Ritter, J. O. (2009). Midlife crisis: A debate. *Gerontology, 55*(5), 582–591. https://doi.org/10.1159/000227322

Fujita, F., & Diener, E. (2005). Life satisfaction set point: Stability and change. *Journal of Personality and Social Psychology, 88*(1), 158–164. https://doi.org/https://doi.org/10.1037/0022-3514.88.1.158

Fung, H. H., Carstensen, L. L., & Lutz, A. M. (1999). Influence of time on social preferences: Implications for life-span development. *Psychology and Aging, 14*(4), 595–604. https://doi.org/10.1037/0882-7974.14.4.595

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

*Gao, M., & Cummings, E. M. (2019). Understanding parent-child relationship as a developmental process: Fluctuations across days and changes over years. *Developmental Psychology, 55*(5), 1046–1058. https://doi.org/10.1037/dev0000680

*Girme, Y. U., Overall, N. C., & Faingataa, S. (2014). "Date nights" take two: The maintenance function of shared relationship activities. *Personal Relationships, 21*(1), 125–149. https://doi.org/10.1111/pere.12020

Gloor, S., Gonin-Spahni, S., Znoj, H., & Perrig-Chiello, P. (2021). Repartnering and trajectories of life satisfaction after separation and divorce in middle and later life. *Journal of Social and Personal Relationships*. https://doi.org/10.1177/02654075211009594 (Advance online publication)

Gottman, J. M., & Levenson, R. W. (1999). Rebound from marital conflict and divorce prediction. *Family Process, 38*(3), 287–292. https://doi.org/10.1111/j.1545-5300.1999.00287.x

*Gray, J. S., & Ozer, D. J. (2019). Comparing two models of dyadic change: Correlated growth versus common fate. *Social Psychological and Personality Science, 10*(7), 957–965. https://doi.org/10.1177/1948550618799066

*Greving Mehall, K., Spinrad, T. L., Eisenberg, N., & Gaertner, B. M. (2009). Examining the relations of infant temperament and couples' marital satisfaction to mother and father involvement: A longitudinal study. *Fathering, 7*(1), 23–48. https://doi.org/10.3149/fth.0701.23

*Gustavson, K., Roysamb, E., Borren, I., Torvik, F. A., & Karevold, E. (2016). Life satisfaction in close relationships: Findings from a longitudinal study [Peer Reviewed]. *Journal of Happiness Studies: An Interdisciplinary Forum on Subjective Well-Being, .17*(3), pp. https://doi.org/10.1007/s10902-015-9643-7

*Hagemeyer, B., Neberich, W., Asendorpf, J. B., & Neyer, F. J. (2013). (In)congruence of implicit and explicit communal motives predicts the quality and stability of couple relationships. *Journal of Personality, 81*(4), 390–402. https://doi.org/10.1111/jopy.12016

*Hakanen, J. J., Peeters, M. C. W., & Perhoniemi, R. (2011). Enrichment processes and gain spirals at work and at home: A 3-year cross-lagged panel study. *Journal of Occupational and Organizational Psychology, 84*(1), 8–30. https://doi.org/10.1111/j.2044-8325.2010.02014.x

Halpern-Meekin, S., Manning, W. D., Giordano, P. C., & Longmore, M. A. (2013). Relationship churning in emerging adulthood: On/off relationships and sex with an ex. *Journal of Adolescent Research, 28*(2), 166–188. https://doi.org/10.1177/0743558412464524

*Hammond, M. D., & Overall, N. C. (2014). Endorsing benevolent sexism magnifies willingness to dissolve relationships when facing partner-ideal discrepancies. *Personal Relationships, 21*(2), 272–287. https://doi.org/10.1111/pere.12031

Havighurst, R. J. (1972). *Developmental tasks and education (3rd ed.)*. McKay.

Heckhausen, J., Dixon, R. A., & Baltes, P. B. (1989). Gains and losses in development throughout adulthood as perceived by different adult age groups. *Developmental Psychology, 25*(1), 109–121. https://doi.org/10.1037/0012-1649.25.1.109

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Academic Press.

Hendrick, S. S. (1988). A generic measure of relationship satisfaction. *Journal of Marriage and the Family, 50*(1), 93–98. https://doi.org/10.2307/352430

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences, 33*(2–3), 61–83. https://doi.org/10.1017/s0140525x0999152x

*Hernandez-Kane, K. M., & Mahoney, A. (2018). Sex through a sacred lens: Longitudinal effects of sanctification of marital sexuality. *Journal of Family Psychology, 32*(4), 425–434. https://doi.org/10.1037/fam0000392

Heuveline, P., & Timberlake, J. M. (2004). The role of cohabitation in family formation: The United States in comparative perspective. *Journal of Marriage and Family, 66*(5), 1214–1230. https://doi.org/10.1111/j.0022-2445.2004.00088.x

Hill, R. (1958). Generic features of families under stress. *Social Casework, 39*(2–3), 139–150. https://doi.org/10.1177/1044389458039002-318

Hogan, R., & Roberts, B. W. (2000). A socioanalytic perspective on person-environment interaction. In W. B. Walsh, K. H. Craik, & H. Price (Eds.), *Person-environment psychology: New directions and perspectives* (2nd ed., pp. 1–23). Lawrence Erlbaum Associates Publishers.

*Hsiao, Y.-L. (2017). Longitudinal changes in marital satisfaction during middle age in Taiwan. *Asian Journal of Social Psychology, 20*(1), 22–32. https://doi.org/10.1111/ajsp.12161

Hudson, N. W., Lucas, R. E., & Donnellan, M. B. (2019). Healthier and happier? A 3-year longitudinal investigation of the prospective associations and concurrent changes in health and experiential well-being [Peer Reviewed]. *Personality and Social Psychology Bulletin, .45*(12), pp. https://doi.org/10.1177/0146167219838547 30975027

Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis: Correcting error and bias in research findings*. Sage.

Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis: Correcting error and bias in research findings* (2nd ed.). Sage.

Hunter, J. E., & Schmidt, F. L. (2014). *Methods of meta-analysis: Correcting error and bias in research findings* (3rd ed.). Sage.

Huston, T. L., Caughlin, J. P., Houts, R. M., Smith, S. E., & George, L. J. (2001). The connubial crucible: Newlywed years as predictors of marital delight, distress, and divorce. *Journal of Personality and Social Psychology, 80*(2), 237–252. https://doi.org/10.1037/0022-3514.80.2.237

Huston, T. L., & Houts, R. M. (1998). The psychological infrastructure of courtship and marriage: The role of personality and compatibility in romantic relationships. In T. N. Bradbury (Ed.), *The developmental course of marital dysfunction* (pp. 114–151). Cambridge University Press. https://doi.org/10.1017/CBO9780511527814.006

Hutteman, R., Hennecke, M., Orth, U., Reitz, A. K., & Specht, J. (2014). Developmental tasks as a framework to study personality development in adulthood and old age. *European Journal of Personality, 28*(3), 267–278. https://doi.org/10.1002/per.1959

Ickes, W., Snyder, M., & Garcia, S. (1997). Personality influences on the choice of situations. In R. Hogan, J. Johnson, & S. Briggs (Eds.), *Handbook of personality psychology* (pp. 166–198). Academic Press.

*Impett, E. A., Kogan, A., English, T., John, O., Oveis, C., Gordon, A. M., & Keltner, D. (2012). Suppression sours sacrifice: Emotional and relational costs of suppressing emotions in romantic relationships. *Personality and Social Psychology Bulletin, 38*(6), 707–720. https://doi.org/10.1177/0146167212437249

*Ivanova, K. (2016). Relationship satisfaction of the previously married: The significance of relationship specific costs and rewards in first and repartnering unions. *Journal of Social and Personal Relationships, 33*(5), 559–580. https://doi.org/10.1177/0265407515583942

*Jayamaha, S. D., & Overall, N. C. (2015). Agents' self-esteem moderates the effectiveness of negative-direct partner regulation strategies. *Personal Relationships, 22*(4), 738–761. https://doi.org/10.1111/pere.12108

*Jenkins, A. I., Fredman, S. J., Le, Y., Sun, X., Brick, T. R., Skinner, O. D., & McHale, S. M. (2020). Prospective associations between depressive symptoms and marital satisfaction in Black couples. *Journal of Family Psychology, 34*(1), 12–23. https://doi.org/10.1037/fam0000573

*Jensen, J. F., & Rauer, A. J. (2015a). Marriage work in older couples: Disclosure of marital problems to spouses and friends over time. *Journal of Family Psychology, 29*(5), 732–743. https://doi.org/10.1037/fam0000099

*Jensen, J. F., & Rauer, A. J. (2015b). Young adult females' relationship work and its links to romantic functioning and stability over time. *Journal of Social and Personal Relationships, 33*(5), 687–708. https://doi.org/10.1177/0265407515588221

*Johnson, M. D., & Anderson, J. R. (2013). The longitudinal association of marital confidence, time spent together, and marital satisfaction. *Family Process, 52*(2), 244–256. https://doi.org/10.1111/j.1545-5300.2012.01417.x

Johnson, M. D., Lavner, J. A., Mund, M., Zemp, M., Stanley, S. M., Neyer, F. J., Impett, E. A., Rhoades, G. K., Bodenmann, G., Weidmann, R., Bühler, J. L., Burriss, R. P., Wünsche, J., & Grob, A. (2021). Within-couple associations between communication and relationship satisfaction over time. *Personality and Social Psychology Bulletin*, 1–16. https://doi.org/10.1177/01461672211016920

Johnson, M. D., & Neyer, F. J. (2019). (Eventual) stability and change across partnerships. *Journal of Family Psychology, 33*(6), 711–721. https://doi.org/10.1037/fam0000523

*Kanat-Maymon, Y., Sarid, O., Mor, Y., Mirsky, J., & Slonim-Nevo, V. (2016). Gaps between immigrant spouses in host country language proficiency: Longitudinal effects on marital satisfaction. *The Journal of Psychology, 150*(6), 793–808. https://doi.org/10.1080/00223980.2016.1196159

Karney, B. R., & Bradbury, T. N. (1995). Assessing longitudinal change in marriage: An introduction to the analysis of growth curves. *Journal of Marriage and the Family, 57*(4), 1091–1108. https://doi.org/10.2307/353425

Karney, B. R., & Bradbury, T. N. (2020). Research on marital satisfaction and stability in the 2010s: Challenging conventional wisdom. *Journal of Marriage and Family, 82*(1), 100–116. https://doi.org/10.1111/jomf.12635

Kazak, A. E. (2018). Editorial: Journal article reporting standards. *American Psychologist, 73*(1), 1–2. https://doi.org/10.1037/amp0000263

Kenny, D. A., & Zautra, A. (2001). Trait–state models for longitudinal data. In L. M. Collins & A. G. Sayer (Eds.), *New methods for the analysis of change* (pp. 243–263). American Psychological Association. https://doi.org/https://doi.org/10.1037/10409-008

*Kerkhof, P., Finkenauer, C., & Muusses, L. D. (2011). Relational consequences of compulsive Internet use: A longitudinal study among newlyweds. *Human Communication Research, 37*(2), 147–173. https://doi.org/10.1111/j.1468-2958.2010.01397.x

*Kluwer, E. S., & Johnson, M. D. (2007). Conflict frequency and relationship quality across the transition to parenthood. *Journal of Marriage and Family, 69*(5), 495–517. https://doi.org/10.1111/j.1741-3737.2007.00434.x

*Kouros, C. D. (2011). *Transactional relations between marital functioning, spouses' depressive symptoms and children's adjustment* [Doctoral dissertation, University of Notre Dame]. University of Notre Dame, Hesburgh Libraries. https://curate.nd.edu/usage/79407w64d4h

Kulu, H. (2014). Marriage duration and divorce: The seven-year itch or a lifelong itch? *Demography, 51*(3), 881–893. https://doi.org/10.1007/s13524-013-0278-1

Kurdek, L. A. (1998). The nature and predictors of the trajectory of change in marital quality over the first 4 years of marriage for first-married husbands and wives. *Journal of Family Psychology, 12*(4), 494–510. https://doi.org/10.1037/0893-3200.12.4.494

Kurdek, L. A. (1999). The nature and predictors of the trajectory of change in marital quality for husbands and wives over the first 10 years of marriage. *Developmental Psychology, 35*(5), 1283–1296. https://doi.org/10.1037/0012-1649.35.5.1283

Kuster, F., & Orth, U. (2013). The long-term stability of self-esteem: Its time-dependent decay and nonzero asymptote. *Personality and Social Psychology Bulletin, 39*(5), 677–690. https://doi.org/10.1177/0146167213480189

*Lavner, J. A., & Bradbury, T. N. (2010). Patterns of change in marital satisfaction over the newlywed years. *Journal of Marriage and Family, 72*(5), 1171–1187. https://doi.org/10.1111/j.1741-3737.2010.00757.x

Lavner, J. A., Bradbury, T. N., & Karney, B. R. (2012). Incremental change or initial differences? Testing two models of marital deterioration. *Journal of Family Psychology, 26*(4), 606–616. https://doi.org/10.1037/a0029052

*LeBaron, C. D. L., Miller, R. B., & Yorgason, J. B. (2014). A longitudinal examination of women's perceptions of marital power and marital happiness in midlife marriages. *Journal of Couple & Relationship Therapy, 13*(2), 93–113. https://doi.org/10.1080/15332691.2013.852492

Levinson, D. J., Darrow, C. M., Klein, E. B., Levinson, M. H., & McKee, B. (1976). Periods in the adult development of men: Ages 18 to 45. *The Counseling Psychologist, 6*(1), 21–25. https://doi.org/10.1177/001100007600600105

Lewis, M. (1997). *Altering fate: Why the past does not predict the future*. Guilford Press.

Lewis, M. (1999). On the development of personality. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (pp. 327–346). Guilford Press.

Lewis, M. (2000a). Issues in the study of personality development. *Psychological Inquiry, 12*(2), 67–83. https://doi.org/10.1207/S15327965PLI1202_02

Lewis, M. (2000b). Models of development. In D. Cervone & W. Mischel (Eds.), *Advances in personality science* (pp. 153–176). Guilford Press.

Li, T., & Chan, D. K. S. (2012). How anxious and avoidant attachment affect romantic relationship quality differently: A meta-analytic review. *European Journal of Social Psychology, 42*(4), 406–419. https://doi.org/https://doi.org/10.1002/ejsp.1842

*Li, X., Cao, H., Zhou, N., Ju, X., Lan, J., Zhu, Q., & Fang, X. (2018). Daily communication, conflict resolution, and marital quality in Chinese marriage: A three-wave, cross-lagged analysis. *Journal of Family Psychology, 32*(6), 733–742. https://doi.org/10.1037/fam0000430

*Lickenbrock, D. M., & Braungart-Rieker, J. M. (2015). Examining antecedents of infant attachment security with mothers and fathers: An ecological systems perspective. *Infant Behavior and Development, 39*, 173–187. https://doi.org/10.1016/j.infbeh.2015.03.003

Light, R. J., & Pillemer, D. B. (1984). *Summing up: The science of reviewing research*. Harvard University Press.

*Lin, W.-F., Li, T.-S., & Chen, L. H. (2017). Sacrifice in a supportive marriage: Social support as a moderator buffers the negative effects of sacrifice in marriage. *Journal of Happiness Studies: An Interdisciplinary Forum on Subjective Well-Being, 18*(2), 575–589. https://doi.org/10.1007/s10902-016-9738-9

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Sage.

Locke, H. J., & Wallace, K. M. (1959, Vol.21). Short marital-adjustment and prediction tests: Their reliability and validity. *Marriage & Family Living, 21*(3), 251–255. https://doi.org/10.2307/348022

*Logan, J. M., & Cobb, R. J. (2013). Trajectories of relationship satisfaction: Independent contributions of capitalization and support perceptions. *Personal Relationships, 20*(2), 277–293. https://doi.org/10.1111/j.1475-6811.2012.01408.x

Lorber, M. F., Erlanger, A. C. E., Heyman, R. E., & O'Leary, K. D. (2015). The honeymoon effect: Does it exist and can it be predicted? *Prevention Science, 16*(4), 550–559. https://doi.org/10.1007/s11121-014-0480-4

Lucas, R. E., & Donnellan, M. B. (2007). How stable is happiness? Using the STARTS model to estimate the stability of life satisfaction. *Journal of Research in Personality, 41*(5), 1091–1098. https://doi.org/https://doi.org/10.1016/j.jrp.2006.11.005

Lykken, D., & Tellegen, A. (1996). Happiness is a stochastic phenomenon [Peer Reviewed]. *Psychological Science, 7*(3), 186–189. https://doi.org/10.1111/j.1467-9280.1996.tb00355.x

McAdams, D. P. (1995). What do we know when we know a person? *Journal of Personality, 63*(3), 365–396. https://doi.org/10.1111/j.1467-6494.1995.tb00500.x

McAdams, D. P. (2013). The psychological self as actor, agent, and author. *Perspectives on Psychological Science, 8*, 272–295. https://doi.org/10.1177/1745691612464657

McAdams, D. P. (2015). Three lines of personality development: A conceptual itinerary. *European Psychologist, 20*(4), 252–264. https://doi.org/10.1027/1016-9040/a000236

McGue, M., Bacon, S., & Lykken, D. T. (1993). Personality stability and change in early adulthood: A behavioral genetic analysis. *Developmental Psychology, 29*(1), 96–109. https://doi.org/10.1037/0012-1649.29.1.96

McNulty, J. K. (2016). Highlighting the contextual nature of interpersonal relationships. *Advances in Experimental Social Psychology, 54*, 247–315. https://doi.org/10.1016/bs.aesp.2016.02.003

McNulty, J. K., Meltzer, A. L., Neff, L. A., & Karney, B. (2021). How both partners' individual differences, stress, and behavior predict change in relationship satisfaction: Extending the VSA model. *Proceedings of the National Academy of Sciences of the United States of America, 118*(27), e2101402118. https://doi.org/10.1073/pnas.2101402118

Mehta, C. M., Arnett, J. J., Palmer, C. G., & Nelson, L. J. (2020). Established adulthood: A new conception of ages 30 to 45. *American Psychologist, 75*(4), 431–444. https://doi.org/10.1037/amp0000600

*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), 418–428. https://doi.org/10.1037/a0034424

*Meltzer, A. L., Novak, S. A., McNulty, J. K., Butler, E. A., & Karney, B. R. (2013). Marital satisfaction predicts weight gain in early marriage. *Health Psychology, 32*(7), 824–827. https://doi.org/10.1037/a0031593

*Menéndez, S., Hidalgo, M. V., Jiménez, L., & Moreno, M. C. (2011). Father involvement and marital relationship during transition to parenthood: Differences between dual and single-earner families. *The Spanish Journal of Psychology, 14*(2), 639–647. https://doi.org/0.5209/rev_SJOP.2011.v14.n2.12

*Miller, P. J. E., Caughlin, J. P., & Huston, T. L. (2003). Trait expressiveness and marital satisfaction: The role of idealization processes. *Journal of Marriage and Family, 65*(4), 978–995. https://doi.org/10.1111/j.1741-3737.2003.00978.x

*Moen, D. A. (2012). *Newlywed to established marriage: A longitudinal study of early risk and protective factors that influence marital satisfaction* (Publication No. 12-2011) [Doctoral dissertation, Utah State University]. Utah State University, Merrill-Cazier Library. https://digitalcommons.usu.edu/etd/1127

Moss, H. A., & Susman, E. J. (1980). Longitudinal study of personality development. In J. O. G. Brim & J. Kagan (Eds.), *Constancy and change in human development* (pp. 530–595). Harvard University Press.

*Mund, M., Finn, C., Hagemeyer, B., Zimmermann, J., & Neyer, F. J. (2015). The dynamics of self-esteem in partner relationships. *European Journal of Personality, 29*(2), 235–249. https://doi.org/10.1002/per.1984

Mund, M., & Neyer, F. J. (2014). Treating personality-relationship transactions with respect: Narrow facets, advanced models, and extended time frames. *Journal of Personality and Social Psychology, 107*(2), 352–368. https://doi.org/10.1037/a0036719

Murray, S. L., Holmes, J. G., & Griffin, D. W. (2000). Self-esteem and the quest for felt security: How perceived regard regulates attachment processes. *Journal of Personality and Social Psychology, 78*(3), 478–498. https://doi.org/0.1037//0022-3514.78.3.478

Murray, S. L., Holmes, J. G., & Griffin, D. W. (2000). Self-esteem and the quest for felt security: How perceived regard regulates attachment processes. *Journal of Personality and Social Psychology, 78*(3), 478-498. https://doi.org/10.1037//0022-3514.78.3.478

*Naud, C., Lussier, Y., Sabourin, S., Normandin, L., Clarkin, J. F., & Kernberg, O. F. (2013). How attachment and excessive self-sacrificing depressive dynamics are related to couple relationship satisfaction over time. *Couple and Family Psychology: Research and Practice, 2*(1), 14–33. https://doi.org/10.1037/a0031705

Neisser, U., Boodoo, G., Bouchard, T. J., Jr., Boykin, A. W., Brody, N., Ceci, S. J., Halpern, D. F., Loehlin, J. C., Perloff, R., Sternberg, R. J., & Urbina, S. (1996). Intelligence: Knowns and unknowns. *American Psychologist, 51*(2), 77–101. https://doi.org/10.1037/0003-066x.51.2.77

Neugarten, B. L. (1968). The awareness of middle age. In B. L. Neugarten (Ed.), *Middle age and aging* (pp. 93–98). University of Chicago Press.

Neyer, F. J., & Asendorpf, J. B. (2001). Personality-relationship transaction in young adulthood. *Journal of Personality and Social Psychology, 81*(6), 1190–1204. https://doi.org/10.1037/0022-3514.81.6.1190

Neyer, F. J., Mund, M., Zimmermann, J., & Wrzus, C. (2014). Personality-relationship transactions revisited. *Journal of Personality, 82*(6), 539–550. https://doi.org/10.1111/jopy.12063

*Nguyen, T. P., Williamson, H. C., Karney, B. R., & Bradbury, T. N. (2017). Communication moderates effects of residential mobility on relationship quality among ethnically diverse couples. *Journal of Family Psychology, 31*(6), 753–764. https://doi.org/10.1037/fam0000324

*Niessen, C., Müller, T., Hommelhoff, S., & Westman, M. (2018). The impact of preventive coping on business travelers' work and private life. *Journal of Organizational Behavior, 39*(1), 113–127. https://doi.org/10.1002/job.2215

Ogilvie, D. M., Rose, K. M., & Heppen, J. B. (2001). A comparison of personal project motives in three age groups. *Basic and Applied Social Psychology, 23*(3), 207–215. https://doi.org/10.1207/s15324834basp2303_7

*Ogolsky, B. G., Surra, C. A., & Monk, J. K. (2016). Pathways of commitment to wed: The development and dissolution of romantic relationships. *Journal of Marriage and Family, 78*(2), 293–310. https://doi.org/10.1111/jomf.12260

Orengo-Aguayo, R. E. (2015). Mexican American and other Hispanic couples' relationship dynamics: A review to inform interventions aimed at promoting healthy relationships. *Marriage & Family Review, 51*(7), 633–667. https://doi.org/https://doi.org/10.1080/01494929.2015.1068253

Orth, U., Erol, R. Y., & Luciano, E. C. (2018). Development of self-esteem from age 4 to 94 years: A meta-analysis of longitudinal studies. *Psychological Bulletin, 144*(10), 1045–1080. https://doi.org/10.1037/bul0000161

*Orth, U., Maes, J., & Schmitt, M. (2015). Self-esteem development across the life span: A longitudinal study with a large sample from Germany. *Developmental Psychology, 51*(2), 248–259. https://doi.org/10.1037/a0038481

*Orth, U., Robins, R. W., & Widaman, K. F. (2012). Life-span development of self-esteem and its effects on important life outcomes. *Journal of Personality and Social Psychology, 102*(6), 1271–1288. https://doi.org/10.1037/a0025558

*Padilla, J., McHale, S. M., De Jesus, S. A. R., Updegraff, K. A., & Umana-Taylor, A. J. (2018). Longitudinal course and correlates of parents' differential treatment of siblings in Mexican-origin families. *Family Process, 57*(4), 979–995. https://doi.org/10.1111/famp.12328

*Paleari, F. G., Regalia, C., & Fincham, F. (2005). Marital quality, forgiveness, empathy, and rumination: A longitudinal analysis. *Personality and Social Psychology Bulletin, 31*(3), 368–378. https://doi.org/10.1177/0146167204271597

Pals, J. L. (1999). Identity consolidation in early adulthood: Relations with ego-resiliency, the context of marriage, and personality change. *Journal of Personality, 67*(2), 295–329. https://doi.org/10.1111/1467-6494.00057

*Parfitt, Y., Ayers, S., Pike, A., Jessop, D., & Ford, E. (2014). A prospective study of the parent–baby bond in men and women 15 months after birth. *Journal of reproductive and infant psychology, 32*(5), 441–456. https://doi.org/10.1080/02646838.2014.956301

*Parise, M., Manzi, C., Donato, S., & Iafrate, R. (2017). Free to love? The role of intrusive parenting for young adult children's romantic relationship quality. *Journal of Prevention & Intervention in the Community, 45*(3), 187–201. https://doi.org/10.1080/10852352.2016.1198127

*Peltz, J. S., Rogge, R. D., & Sturge-Apple, M. L. (2018). Transactions within the family: Coparenting mediates associations between parents' relationship satisfaction and the parent–child relationship. *Journal of Family Psychology, 32*(5), 553–564. https://doi.org/10.1037/fam0000413

Proulx, C. M., Helms, H. M., & Buehler, C. (2007). Marital quality and personal well-being: A meta-analysis. *Journal of Marriage and Family, 69*(3), 576–593. https://doi.org/10.1111/j.1741-3737.2007.00393.x

R Development Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org

Rammstedt, B., & Schupp, J. (2008). Only the congruent survive--Personality similarities in couples. *Personality and Individual Differences, 45*(6), 533–535. https://doi.org/10.1016/j.paid.2008.06.007

Randall, A. K., & Bodenmann, G. (2009). The role of stress on close relationships and marital satisfaction. *Clinical Psychology Review, 29*(2), 105–115. https://doi.org/10.1016/j.cpr.2008.10.004

*Reizer, A., Ein-Dor, T., & Shaver, P. (2014). The avoidance cocoon: Examining the interplay between attachment and caregiving in predicting relationship satisfaction. *European Journal of Social Psychology, 44*(7), 774–786. https://doi.org/10.1002/ejsp.2057

*Roberson, P. N., Fish, J. N., Olmstead, S. B., & Fincham, F. D. (2015). College adjustment, relationship satisfaction, and conflict management: A cross-lag assessment of developmental "spillover". *Emerging Adulthood, 3*(4), 244-254. https://doi.org/10.1177/2167696815570710

Roberts, B. W., & Caspi, A. (2003). The cumulative continuity model of personality development: Striking a balance between continuity and change in personality traits across the life course. In R. M. Staudinger & U. Lindenberger (Eds.), *Understanding human development: Lifespan psychology in exchange with other disciplines*. Kluwer Academic.

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*, 3–25. https://doi.org/10.1037/0033-2909.126.1.3

Roberts, B. W., & Nickel, L. B. (2021). Personality development across the life course: A neo-socioanalytic perspective. In O. P. John & R. W. Robins (Eds.), *Handbook of personality: Theory and research* (pp. 259-283). Guilford.

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), 1–25. https://doi.org/10.1037/0033-2909.132.1.1

Roberts, B. W., & Wood, D. (2006). Personality development in the context of the neo-socioanalytic model of personality. In D. K. Mroczek & T. D. Little (Eds.), *Handbook of personality development* (pp. 11–39). Erlbaum.

*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 of Personality, 70*(6), 925–964. https://doi.org/10.1111/1467-6494.05028

Robins, R. W., Fraley, R. C., Roberts, B. W., & Trzesniewski, K. H. (2001). A longitudinal study of personality change in young adulthood. *Journal of Personality, 69*(4), 617–640. https://doi.org/https://doi.org/10.1111/1467-6494.694157

Robins, R. W., Hendin, H. M., & Trzesniewski, K. H. (2001). Measuring global self-esteem: Construct validation of a single-item measure and the Rosenberg Self-Esteem Scale. *Personality and Social Psychology Bulletin, 27*(2), 151–161. https://doi.org/0146167201272002

Robles, T. F., Slatcher, R. B., Trombello, J. M., & McGinn, M. M. (2014). Marital quality and health: A meta-analytic review. *Psychological Bulletin, 140*(1), 140–187. https://doi.org/10.1037/a0031859

Rothstein, H. R., Sutton, A. J., & Borenstein, M. (2005). *Publication bias in meta-analysis: Prevention, assessment and adjustments*. John Wiley & Sons. https://doi.org/10.1002/0470870168.ch1

*Ruffieux, M., Nussbeck, F. W., & Bodenmann, G. (2014). Long-term prediction of relationship satisfaction and stability by stress, coping, communication, and well-being. *Journal of Divorce & Remarriage, 55*(6), 485–501. https://doi.org/10.1080/10502556.2014.931767

Rusbult, C. E. (1980). Commitment and satisfaction in romantic associations: A test of the investment model. *Journal of Experimental Social Psychology, 16*(2), 172–186. https://doi.org/10.1016/0022-1031(80)90007-4

Rusbult, C. E. (1983). A longitudinal test of the investment model: The development (and deterioration) of satisfaction and commitment in heterosexual involvements. *Journal of Personality and Social Psychology, 45*(1), 101–117. https://doi.org/10.1037/0022-3514.45.1.101

Rusbult, C. E., Martz, J. M., & Agnew, C. R. (1998). The Investment Model Scale: Measuring commitment level, satisfaction level, quality of alternatives, and investment size. *Personal Relationships, 5*(4), 357–387. https://doi.org/10.1111/j.1475-6811.1998.tb00177.x

*Sadikaj, G., Rappaport, L. M., Moskowitz, D., Zuroff, D. C., Koestner, R., & Powers, T. (2015). Consequences of interpersonal spin on couple-relevant goal progress and relationship satisfaction in romantic relationships. *Journal of Personality and Social Psychology, 109*(4), 722–737. https://doi.org/10.1037/pspp0000043

Sanchez, L., & Thomson, E. (1997). Becoming mothers and fathers: Parenthood, gender, and the division of labor. *Gender & Society, 11*, 747–772. https://doi.org/10.1177/089124397011006003

Sbarra, D. A., Law, R. W., & Portley, R. M. (2011). Divorce and death: A meta-analysis and research agenda for clinical, social, and health psychology. *Perspectives on Psychological Science, 6*(5), 454–474. https://doi.org/10.1177/1745691611414724

Scharfe, E., & Bartholomew, K. (1994). Reliability and stability of adult attachment patterns. *Personal Relationships, 1*(1), 23–43. https://doi.org/https://doi.org/10.1111/j.1475-6811.1994.tb00053.x

Schimmack, U., & Oishi, S. (2005). The influence of chronically and temporarily accessible information on life satisfaction judgments. *Journal of Personality and Social Psychology, 89*(3), 395–406. https://doi.org/https://doi.org/10.1037/0022-3514.89.3.395

*Schober, P. S. (2012). Paternal child care and relationship quality: A longitudinal analysis of reciprocal associations. *Journal of Marriage and Family, 74*(2), 281–296. https://doi.org/10.1111/j.1741-3737.2011.00955.x

Schröder, J., & Schmiedeberg, C. (2015). Effects of relationship duration, cohabitation, and marriage on the frequency of intercourse in couples: Findings from German panel data. *Social Science Research, 52*, 72–82. https://doi.org/10.1016/j.ssresearch.2015.01.009

Schuerger, J. M., Zarrella, K. L., & Hotz, A. S. (1989). Factors that influence the temporal stability of personality by questionnaire. *Journal of Personality and Social Psychology, 56*(5), 777–783. https://doi.org/https://doi.org/10.1037/0022-3514.56.5.777

Shulman, S., & Connolly, J. (2013). The challenge of romantic relationships in emerging adulthood: Reconceptualization of the field. *Emerging Adulthood, 1*(1), 27–39. https://doi.org/10.1177/2167696812467330

*Sotskova, A., Woodin, E. M., & Gou, L. H. (2015). Hostility, flooding, and relationship satisfaction: Predicting trajectories of psychological aggression across the transition to parenthood. *Aggressive Behavior, 41*(2), 134–148. https://doi.org/10.1002/ab.21570

*South, S. C., Boudreaux, M. J., & Oltmanns, T. F. (2020). The impact of personality disorders on longitudinal change in relationship satisfaction in long-term married couples. *Journal of Personality Disorders, 34*(4), 439-458. https://doi.org/10.1521/pedi_2019_33_392

Spanier, G. B. (1976). Measuring dyadic adjustment: New scales for assessing the quality of marriage and similar dyads. *Journal of Marriage and the Family, 38*(1), 15–28. https://doi.org/10.2307/350547

Specht, J., Egloff, B., & Schmukle, S. C. (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), 862–882. https://doi.org/10.1037/a0024950

Sterne, J. A., & Egger, M. (2001). Funnel plots for detecting bias in meta-analysis: guidelines on choice of axis. *Journal of Clinical Epidemiology, 54*(10), 1046–1055. https://doi.org/10.1016/s0895-4356(01)00377-8

*Sullivan, T. J., Feinstein, B. A., Marshall, A. D., & Mustanski, B. (2017). Trauma exposure, discrimination, and romantic relationship functioning: A longitudinal investigation among LGB young adults. *Psychology of Sexual Orientation and Gender Diversity, 4*(4), 481–490. https://doi.org/10.1037/sgd0000249

*Sun, X., McHale, S. M., Crouter, A. C., & Jones, D. E. (2017). Longitudinal links between work experiences and marital satisfaction in african american dual-earner couples. *Journal of Family Psychology, 31*(8), 1029–1039. https://doi.org/10.1037/fam0000381

Sutton, A. J. (2009). Publication bias. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (pp. 435–452). Russell Sage Foundation.

Swann, W. B., & Read, S. J. (1981). Self-verification processes: How we sustain our self-conceptions. *Journal of Experimental Social Psychology, 17*(4), 351–372. https://doi.org/10.1016/0022-1031(81)90043-3

*Szepsenwol, O., Mizrahi, M., & Birnbaum, G. E. (2015). Fatal suppression: The detrimental effects of sexual and attachment deactivation within emerging romantic relationships. *Social Psychological and Personality Science, 6*(5), 504–512. https://doi.org/0.1177/1948550615572636

Terracciano, A., Costa, P. T., Jr., & McCrae, R. R. (2006). Personality plasticity after age 30. *Personality and Social Psychology Bulletin, 32*(8), 999–1009. https://doi.org/10.1177/0146167206288599

*Tombeau Cost, K., Jonas, W., Unternaehrer, E., Dudin, A., Szatmari, P., Gaudreau, H., Kennedy, J., Atkinson, L., Steiner, M., Meaney, M., & Fleming, A. (2018). Maternal perceptions of paternal investment are associated with relationship satisfaction and breastfeeding duration in humans. *Journal of Family Psychology, 32*(8), 1025–1035. https://doi.org/10.1037/fam0000468

*Tremblay, S., & Pierce, T. (2011). Perceptions of fatherhood: Longitudinal reciprocal associations within the couple. *Canadian Journal of Behavioural Science, 43*(2), 99–110. https://doi.org/10.1037/a0022635

Trzesniewski, K. H., Donnellan, M. B., & Robins, R. W. (2003). Stability of self-esteem across the life span. *Journal of Personality and Social Psychology, 84*(1), 205–220. https://doi.org/https://doi.org/10.1037/0022-3514.84.1.205

*van den Troost, A., Vermulst, A. A., Gerris, J. R., & Matthijs, K. (2005). The Dutch marital satisfaction and communication questionnaire: A validation study. *Psychologica Belgica, 45*(3), 185–206. https://doi.org/10.5334/pb-45-3-185

*van Scheppingen, M. A., Denissen, J. J. A., Chung, J. M., Tambs, K., & Bleidorn, W. (2018). Self-esteem and relationship satisfaction during the transition to motherhood. *Journal of Personality and Social Psychology, 114*(6), 973–991. https://doi.org/10.1037/pspp0000156

Vater, A., & Schröder-Abé, M. (2015). Explaining the link between personality and relationship satisfaction: Emotion regulation and interpersonal behaviour in conflict discussions. *European Journal of Personality, 29*(2), 201–215. https://doi.org/10.1002/per.1993

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software, 36*(3), 1–48. https://doi.org/10.18637/jss.v036.i03

*Volling, B. L., Oh, W., Gonzalez, R., Kuo, P. X., & Yu, T. (2015). Patterns of marital relationship change across the transition from one child to two. *Couple and Family Psychology: Research and Practice, 4*(3), 177–197. https://doi.org/10.1037/cfp0000046

Wagner, J., Ludtke, O., & Trautwein, U. (2016). Self-esteem is mostly stable across young adulthood: Evidence from latent STARTS models. *Journal of Personality, 84*(4), 523−535. https://doi.org/10.1111/jopy.12178

*Weidmann, R., Ledermann, T., & Grob, A. (2017). The interdependence of personality and satisfaction in couples. *European Psychologist, 21*(4), 284–295. https://doi.org/10.1027/1016-9040/a000261

*Weidmann, R., Schönbrodt, F. D., Ledermann, T., & Grob, A. (2017). Concurrent and longitudinal dyadic polynomial regression analyses of Big Five traits and relationship satisfaction: Does similarity matter? *Journal of Research in Personality, 70*, 6–15. https://doi.org/10.1016/j.jrp.2017.04.003

*Whiteman, S. D., McHale, S. M., & Crouter, A. C. (2007). Longitudinal changes in marital relationships: The role of offspring's pubertal development. *Journal of Marriage and Family, 69*(4), 1005–1020. https://doi.org/10.1111/j.1741-3737.2007.00427.x

*Whitton, S. W., Rhoades, G. K., & Whisman, M. A. (2014). Fluctuation in relationship quality over time and individual well-being: Main, mediated, and moderated effects. *Personality and Social Psychology Bulletin, 40*(7), 858–871. https://doi.org/10.1177/0146167214528988

Woods, S. A., & Hampson, S. E. (2005). Measuring the Big Five with single items using a bipolar response scale. *European Journal of Personality, 19*(5), 373–390. https://doi.org/10.1002/per.542

*Zee, K. S., & Weiss, D. (2019). High-quality relationships strengthen the benefits of a younger subjective age across adulthood. *Psychology and Aging, 34*(3), 374–388. https://doi.org/10.1037/pag0000349

Supplemental Material

Table S1

Estimates of Rank-Order Stability of Relationship Satisfaction as a Function of Relationship Duration (Sensitivity Analyses for Using 2 and 6 Years as Estimates of the Average Difference Between Relationship Duration and Duration of Living Together, and Relationship Duration and Marriage Duration, Respectively)

Relationship-duration group	*k*	*N*	Weighted mean effect size	95% CI	*Q*	Variances	
						σ_1^2	σ_2^2
Without time lag as covariate							
0–5 years	37	5,049	**.652**	[.578, .715]	**414.9**	.042	.052
5–10 years	114	22,982	**.736**	[.663, .796]	**6094.5**	.073	.167
10–20 years	98	19,499	**.812**	[.742, .864]	**8914.5**	.243	.092
20–46 years	63	8,071	**.885**	[.809, .932]	**17297.8**	.335	.164
With time lag as covariate							
0–5 years	37	5,049	**.583**	[.456, .687]	**362.5**	.043	.047
5–10 years	114	22,982	**.695**	[.609, .764]	**5615.8**	.094	.130
10–20 years	98	19,499	**.821**	[.761, .867]	**5604.9**	.194	.089
20–46 years	63	8,071	**.904**	[.835, .945]	**9661.1**	.363	.143

Note. Computations without control of time lag were made with multilevel random-effects models and computations with control of time lag were made with multilevel mixed-effects meta-regression models. In the analyses with time lag as covariate, time lag was grand-mean centered at 2.30 years. *k* = number of effect sizes; *N* = number of participants in each group. Weighted mean effect size = disattenuated correlation coefficient, indicating the rank-order stability of relationship satisfaction between two assessments. CI = confidence interval; *Q* = test statistic of the test for (residual) heterogeneity; σ_1^2 = variance proportion attributable to the level of the grouping variable (i.e., between samples); σ_2^2 = variance proportion attributable to the level nested within the grouping variable (i.e., within samples). Values in bold are significant at $p < .05$.

Table S2

Overview of Measures in the Meta-Analytic Data Set and Their Categorization Into Global Satisfaction Measures and Adjustment Measures

Global satisfaction measure
Couples Satisfaction Index (CSI; Funk & Rogge, 2007)
Global Measure of Relationship Satisfaction (GMRS; Lawrance & Byers, 1998)
Intimate Relations Questionnaire (RRQ Braiker & Kelley, 1979)
Kansas Marital Satisfaction Scale (KMSS; Schumm et al., 1986)
Kansas Marital Satisfaction Scale–Revised (KMSS-R; Chung, 2004)
Marital Opinion Questionnaire (MOQ; Huston & Vangelisti 1991)
Relationship Assessment Scale (RAS; Hendrick, 1988)
Relationship satisfaction subscale from the Investment Model Scale (RSI; Rusbult, Martz, & Agnew, 1998)
Self-Report Relationship Satisfaction (SRS; Schmitt et al., 1997)
Quality of Marriage Index (QMI; Norton, 1983)
Adjustment measure
Couple Relationship Domains Questionnaire (CRDQ; Huston, McHale, & Crouter, 1986)
Dyadic Adjustment Scale (DAS; Spanier, 1976)
ENRICH Marital Satisfaction (ENRICH; Fowers & Olson, 1993)
Marital Relationship Inventory (MRI; Burgess, Locke, & Thomes, 1971)
Marital Adjustment Test (MAT; Locke & Wallace, 1959)
Marital Satisfaction Questionnaire for Older Persons (MSQO; Haynes et al., 1992)
Partnership Questionnaire (PFB; Partnerschaftsfragebogen; Hahlweg, 1996)
Perceived Relationship Quality Components Inventory (PQRC; Fletcher, Simpson, & Thomas, 2000), mean
Relationship Evaluation Questionnaire (REQ; Busby et al., 2001)
Relationship Satisfaction Scale (RSS; Gilford & Bengtson, 197)

Table S3

Descriptive Information on Time Lag

Group	*k*	*M*	*SD*	Range
Age				
20–30 years	157	1.88	2.56	0.17–15.00
30–40 years	128	1.95	3.28	0.17–20.00
40–50 years	66	4.24	3.92	0.50–15.00
50–71 years	45	2.08	1.84	0.48–9.00
Relationship duration				
0–5 years	79	1.36	0.97	0.17–5.00
5–10 years	99	1.37	1.75	0.17–15.00
10–20 years	79	3.30	4.65	0.17–20.00
20–46 years	55	3.13	3.26	0.48–15.00

Note. Time lag is given in years. *k* = number of effect sizes

Table S4

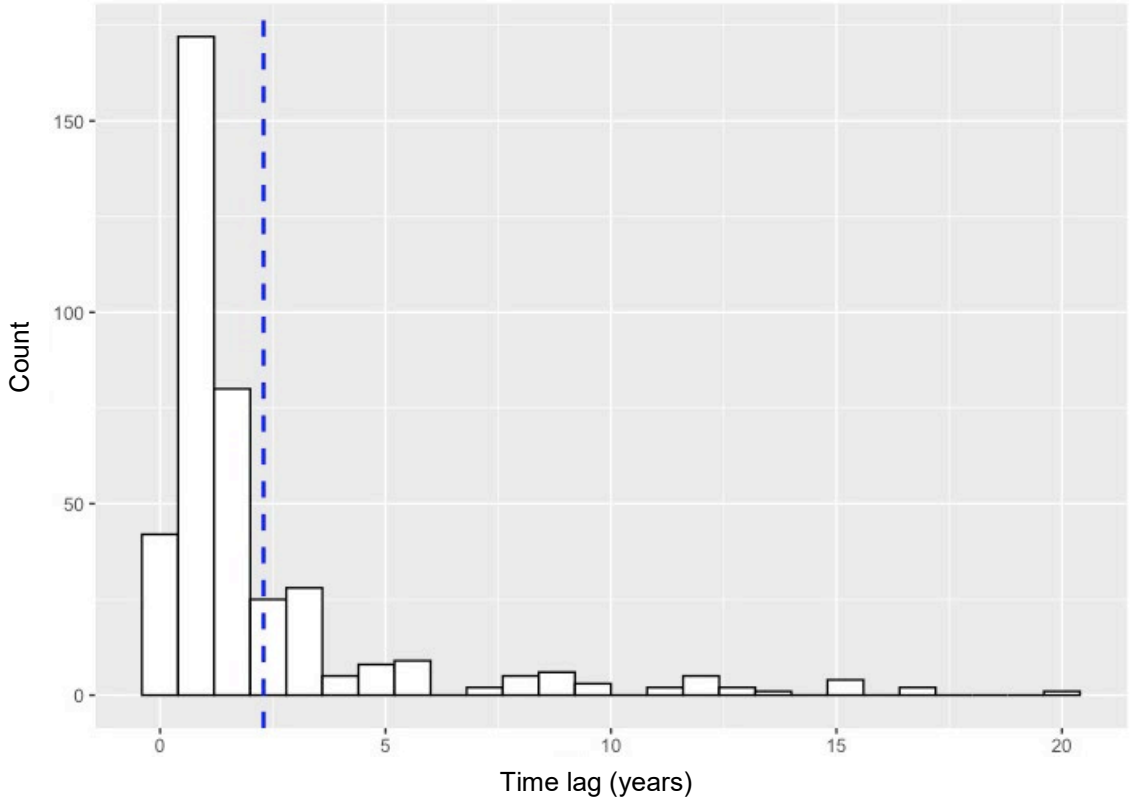
Zero-Order Correlations Between Age, Relationship Duration, Sample Characteristics, Methodological Characteristics, and Effect Size Measure

Variable	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*
1. Age	—																
2. Relationship duration	**.97**	—															
3. Year of birth	**−.72**	**−.75**	—														
4. Sample type	**.12**	**.25**	**−.30**	—													
5. Country	**.12**	**.22**	**−.28**	**−.15**	—												
6. Ethnicity	**.37**	**.42**	**−.21**	**.25**	**−.18**	—											
7. Female	−.09	.01	.07	.09	−.05	.02											
8. Shared household	**.47**	**.43**	**−.41**	**.17**	**.42**	**.31**	−.02	—									
9. Married	**.26**	**.30**	**−.38**	.01	**.33**	.08	.00	**.86**	—								
10. Children	**.70**	**.64**	−.03	.06	.12	**.36**	.03	**.28**	.08	—							
11. Transition	**−.16**	**−.18**	**.26**	−.05	**−.10**	**.21**	.00	.06	**−.11**	−.09	—						
12. Post transition	**−.27**	**−.24**	.05	**−.12**	**.11**	−.06	−.02	.08	**.13**	**−.15**	**−.14**	—					
13. Baseline mean	**−.42**	**−.29**	**.28**	**.25**	**−.20**	−.10	.00	**−.38**	−.05	**−.51**	**.12**	**.20**	—				
14. Time lag	**.12**	**.25**	**−.48**	**.43**	.07	**.17**	.03	.14	**.12**	.03	**−.15**	.00	.08	—			
15. Dyadic data	−.10	−.07	**.22**	**−.43**	**.21**	**−.19**	**−.16**	.09	.03	**−.15**	**.12**	.08	.00	**−.35**	—		
16. Measure (established)	−.05	**−.12**	**.19**	**−.68**	.09	−.04	−.09	−.13	**−.12**	−.09	.06	.02	**−.32**	**−.38**	**.38**	—	
17. Measure (global)	**−.28**	**−.28**	**.39**	**.16**	**−.20**	**.22**	.03	.04	**−.21**	**.19**	**.26**	−.02	**.47**	.01	**−.14**	n.a.	—
18. Effect size	**.32**	**.29**	**−.10**	**−.22**	**.28**	−.02	−.06	**.20**	**.22**	**.16**	.05	**−.14**	**−.35**	**−.36**	**.18**	**.12**	**−.37**

Note. Effect size = disattenuated correlation coefficient, indicating the rank-order stability of relationship satisfaction between Time 1 and Time 2. The following variables were dichotomous: sample type (1 = community or student sample, 0 = nationally representative sample), country, (1 = USA, 0 = other), ethnicity (1 = White, 0 = other), transition (1 = relationship transition between Time 1 and Time 2, 0 = no relationship transition between Time 1 and Time 2), post transition (1 = relationship transition shortly before Time 1 , 0 = no relationship transition shortly before Time 1), dyadic data (1 = dyadic data, 0 = no dyadic data), the two contrasts for measure of relationship satisfaction (1 = established scale, 0 = ad-hoc measure; 1 = global satisfaction, 0 = adjustment measure). Baseline mean refers to mean of relationship satisfaction at Time 1 in the metric of POMP scores, and time lag refers to the interval between Time 1 and Time 2. n.a. indicates that this correlation could not be calculated because of the nested structure of the variables. Values in bold are significant at $p < .05$.

Figure S1

Frequency Distribution of Time Lag Across Effect Sizes



Note. Time lag refers to the time interval between two assessments. The vertical blue line represents the average time lag (2.30 years).