2022-09-14

Neuroticism and Relationship Quality: A Meta-Analytic Review

Charlotte R. Esplin
Brigham Young University

Neuroticism and Relationship Quality: A Meta-Analytic Review

Charlotte R. Esplin

A dissertation submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Scott Braithwaite, Chair
Patrick Steffen
Alan Hawkins
Wendy Birmingham

Department of Psychology

Brigham Young University

ABSTRACT

Neuroticism and Relationship Quality: A Meta-Analytic Review

Charlotte R. Esplin
Department of Psychology, BYU
Doctor of Philosophy

Of each of the Big-Five personality traits, neuroticism is most strongly associated with poor relationship outcomes. Prior research has established a robust negative association between neuroticism and relationship quality, regardless of participant age, length of relationship, and country of origin. As so much has already been studied on the topic of neuroticism and relationship quality, the current study explores the association between neuroticism and relationship using meta-analytic methodology, and details whether Hill's seven criteria for causation have been met in the current literature thus far. After searching through databases and the references of included studies, I used 151 published studies that reported an effect size between neuroticism and relationship quality. Reported effect sizes included an overall aggregate correlation of -.222, and differentiated effect sizes for male actor and partner correlations, and female actor and partner correlations. I used two-way tests to explore if the sample country of origin, type of measure used, and whether the effect size was cross-sectional or longitudinal moderated the -.222 association. Further, I used meta-regression to test whether the length of relationship or participant age moderated the association. Results showed that sample country of origin and participant age did not moderate the association, suggesting that neuroticism and relationship quality are similarly associated regardless of where a participant was from or their age. However, longitudinal data showed a stronger negative relationship than cross-sectional data, measures of depression and anxiety demonstrated a stronger negative relationship than broad neuroticism measures, and a greater time spent in a relationship showed a weaker relationship between the two variables. Results are discussed in light of the interpersonal and intrapersonal models of neuroticism and the social exchange theory, and a model for how neuroticism operates through emotions, interpretations and behaviors is outlined. This model offers information for ways a couples' therapist could work to mitigate the association between neuroticism and relationship quality. A case for causation could not be established, so future directions for the field are delineated.

Keywords: relationships, satisfaction, neuroticism, personality variables

ACKNOWLEDGEMENTS

To my mentor, who has opened doors that I did not know existed;
to my parents, who told me I could do anything I put my mind to;
and to my husband, who has always given me the wings to fly.

TABLE OF CONTENTS

Title Page i

Abstract ii

Acknowledgements iii

List of Figures vi

Neuroticism and Relationship Quality: A Meta-Analytic Review 1

Literature Review: History of Personality Psychology 2

A Closer Look at Neuroticism 9

How is Neuroticism Measured? 11

Relationship Quality 14

How Does Neuroticism Affect Relationship Quality? 17

A Case for Causation 21

The Current Study 22

Method 23

Data Storage and Management 23

Inclusion and Exclusion Criteria 24

Coding 25

Moderators and Study Characteristics 26

Eliminating Bias 28

Power 28

Data Analysis 29

Results 30

Overall Effect Sizes 30

Moderator Analyses 31

Case for Causation 35

Discussion 36

Can Causation Be Established? 40

Future Research 42

Theoretical Underpinnings 43

How Does Neuroticism Operate in Relationships? 46

Clinical Applications 48

Strengths and Limitations 49

References 52

Appendix A 71

Appendix B 76

LIST OF FIGURES

Figure 1 *Flowchart for Inclusion/Exclusion Decisions* 25

Neuroticism and Relationship Quality: A Meta-Analytic Review

Although most married couples vow to part only at death, approximately 42% to 50% of marriages end in divorce (Bradbury et al., 2000). Recently, some have called divorce a public health concern, as children of divorced parents experience more mental health problems (OR = 1.56; 95% CI [1.31, 1.86]; Sands et al., 2017), changes in their adrenocortical responses and the synaptic development of the anterior cingulate cortex, and a higher likelihood of marital distress and divorce (Amato, 1999; Auersperg et al., 2019; Chun et al., 2016; Vezzetti, 2016). On the contrary, those in happy marriages engage in less crime, live longer lives, and experience less psychopathology (Laub et al., 2017; Stack & Eshleman, 1998; Waite & Gallagher, 2001). For these reasons, relationship quality research has proliferated (Bradbury et al., 2000). The current rate of divorce can be attributed to many factors: replicating parental marital instability (Wolfinger, 2000), infidelity (Diefenbach & Opp, 2007), weakened commitment to one another (Whitton et al., 2013), poorer relationship skills (Stanley et al., 2020) and the personalities of the partners (Solomon & Jackson, 2014).

Personality traits have been shown to account for over 60% of the variance associated with relationship outcomes (Russell & Wells, 1994a). Empirical evidence suggests that neuroticism is most consistently associated with poor relationship outcomes (Costa & McCrae, 1988; Cundiff et al., 2012; Karney & Bradbury, 1997; Russell & Wells, 1994a; Russell & Wells, 1994b). According to Costa and McCrae, neuroticism is a general tendency to experience negative emotions like depression, anxiety, or anger in response to stress or frustration (Costa et al., 1986; Lahey, 2009). In a relationship, neuroticism likely leads to excessive emotionality, poor communication, and more conflict (Iveniuk et al., 2014; Mund et al., 2016; Tong et al., 2018). The current literature strongly suggests that neuroticism decreases relationship quality

(Karney & Bradbury, 1997; Russell & Wells, 1994a), likely via both interpersonal and intrapersonal pathways. In the present study, I review and summarize the literature on neuroticism and relationship quality in the context of the Bradford Hill criteria for causation (1965), explain the shortcomings in the field currently, and present a model for how these two constructs operate in relation to one another.

Literature Review: History of Personality Psychology

No single definition of personality is accepted by all personality theorists (Mischel et al., 2007; Sherman, 1941). That said, one broad definition of personality is an individual's unique variation, adaptation to challenges, and self-defining life narratives (McAdams & Pals, 2006). According to personality theorist Hans Eysenck, personality is "the more or less stable and enduring organization of a person's character, temperament, intellect, and physique, which determines his unique adjustment to the environment" (Eysenck, 1970, p. 2). Generally, personality characteristics are considered to be comprised of dispositions, powers, and derivatives (Bergner, 2020). Dispositions, or the personal inclinations that an individual holds, includes his or her traits, styles, attitudes, and interests (Bergner, 2020). The "powers" of an individual's personality include their abilities, knowledge, beliefs, and values (Bergner, 2020). Lastly, an individual's personality derivatives include states (a fleeting mood or feeling that is not driven by their ordinary disposition or powers, i.e., being enraged, being drunk), capacities (an individual's potential for learning), and embodiment (the person's physical characteristics) (Bergner, 2020). Without a consensus definition of personality, theorists throughout the years have sought to understand what personality is and how it interacts with the environment. The field of personality psychology sought to scientifically study humans as a complex whole and to

understand the interrelationship between human motivation, biological drives, and cognitive processes (Geen et al., 1984).

First emerging in the late 1800s, some of the earliest theories of personality include Freud's theories of the Oedipus complex, penis envy, and how sexuality and aggression comprise most of a human's personality (Caprara, 1992; Dumont, 2010). These theories were based on discrete stages in which erotic stimulation of the mouth, anus, and genital area becomes the focus of the child's libido at different ages, most of which take place in the unconscious (Elkatawneh, 2013). Freud offered an early explanation of neuroticism; namely, that it was derived from either real or imagined infantile sexual abuse (Forrester, 1991). The abused individual would go forward with life, only with added anxieties, insecurities, obsessions, and depression. The key problem with Freud's theories was that they were unfalsifiable, which contributed to how the field of personality psychology lacked the theoretical framework necessary to be considered a robust field (Ricoeur, 1992). Common arguments by critics were that it was too "soft" in its scientific practices and that the role of situational factors in an individual's choices was largely overlooked (Caprara, 1992; McAdams & Pals, 2006). In order to be considered a scientific field, theorists had to address these concerns.

After Freud, many personality psychologists offered their theories about personality in general, and neuroticism in particular, and how it interacted with the environment in shaping personality. Although Karen Horney was largely influenced by Freudian thought, her theories have been described as radically different from contemporary psychoanalysis (Miletic, 2002). Perhaps one such theory is her idea that neuroticism is a result of a cold and unloving home environment, not exclusively from real or imagined childhood sexual abuse (Hall & Lindzey, 1985). Further, Horney believed that to understand personality, one must look at the totality of

the person (i.e., their spirituality, culture, cognition, and physiochemistry) rather than just their childhood or sexual experiences (Horney, 2013).

Similar to Horney, Adler pushed back against Freud's theories in more ways than one. First, he posited that human behavior is shaped more by future goals than by past experiences, like those from childhood (Adler, 1924). Further, he believed that personality was a dynamic, ever-changing entity that could not be laid out in stages like Freud's psychosexual stages of development (Adler, 1924). Adler's view that humans are intellectual and creative but that they operate in an environment that is out of their control presaged the humanistic psychologists to come (Hall & Lindzey, 1985). Though the theorists were beginning to push back against Freud's teachings, it would be a long time before the field settled on an empirically based theory of personality.

Throughout the years, different theorists would touch on similar topics, but all in a slightly different way. Murray developed a theory of personality in which an individual's behaviors are a result of their needs and drives (Billstedt et al., 2017). According to Murray, people act in response to primary needs like hunger, thirst, and sexuality, and secondary needs like nurturance, achievement, and social affiliation (Billstedt et al., 2017; Hall & Lindzey, 1985). Murray's theories led him to coin the term "thema," which is defined as the interaction between a need, described above, and "press." "Press" is considered the terms of the environment, or what the environment can do to help or hinder the subject (Hall & Lindzey, 1985). Murray's work into thema led to the creation of the thematic apperception test, which is a popular and commonly used subjective personality test today (Lundy, 1988). Despite its popularity, there have been some claims that the greater the clinician's prior knowledge about the subject, the more valid the clinician's interpretations of the subject's responses (Murstein, 1963). Such a

finding reflects that it is not the test itself that lends valid information about the subject, but instead, how well the clinician already knows the subject. Thus, the field still had some way to go before its theories and tests were considered scientific.

The field of personality psychology began to include more empirical approaches when Allport also rejected the teachings of Freud (McAdams, 1997). Instead, he argued that conscious thoughts and drives were a lot more important than unconscious needs and drives, and (like Adler before him) that the present and future were more important determinants of behavior than the past (Hall & Lindzey, 1985). To understand personality more fully, Allport scoured a dictionary looking for every term that could possibly describe a personality and compiled a list of 4500 terms, which he organized into three categories: cardinal traits, central traits, and secondary traits (Schultz & Schultz, 2017; Sherman, 1941). Cardinal traits are considered the rarest traits; they are so intrinsic that they are almost synonymous with the individual themselves (i.e., Mother Teresa is known for her charity and kindness, so these would be her cardinal traits; Schultz & Schultz, 2017; Sherman, 1941). Central traits are more common, and are major descriptive terms like quiet, shy, outgoing, or angry (Schultz & Schultz, 2017; Sherman, 1941). Lastly, secondary traits are more state-like and are considered responses to certain situations (Schultz & Schultz, 2017; Sherman, 1941). Allport's contributions shaped the path for the beginning of a more empirically sound field of personality psychology.

Behaviorist theories, like those attributed to B.F. Skinner, posit that behavior is explained by observable antecedents and consequences, which shape a personality (Skinner, 1981). Through studying rats, Skinner believed that behavior could be predicted and controlled, and that all behavior is learned through reinforcement (whether reinforcement increases a behavior or decreases a behavior) (Skinner, 1981). Similarly, Bandura argued that behavior is a result of

observation, imitation, and reinforcement – in such, behavior is largely acquired rather than instinctive (Bandura, 1977). For behaviorists, personality was wholly a product of an individual's environment.

Rejecting both the Freudian notion that humans are guided only by aggression, selfishness, and sexuality, and the behaviorist notion that humans are only comprised of their environments, humanistic theorists like Maslow and Rogers believed that humans were innately good, and that they have the potential for healthy and creative growth (Rogers, 2013). Humanism was the first approach that asserted the overall dignity and worth of the human (Hall & Lindzey, 1985). If social and economic barriers would allow all physiological needs to be met, and if an individual feels loved, secure, and has good self-esteem, he or she would have the ability to self-actualize (that is, reach their idealized version of themselves and fulfill their ultimate potential; Crandall et al., 2019). Maslow argued that if these needs were not met, an individual has a greater likelihood of experiencing neurotic symptoms (Lester et al., 1983). In particular, Maslow asserted that people are defined by peak experiences, in which they become "at one" with the world. After having a peak experience, a common effect is a lessening of neurotic symptoms and a healthier view of oneself (Hall & Lindzey, 1985). At this point, neuroticism had not been coined as one of the Big Five factors of personality, which was to come later.

In 1955, Meehl and Cronbach elucidated the concept of construct validity, which pushed personality psychologists to make their constructs more precise (Cronbach & Meehl, 1955). Prior to this, imprecise constructs had led to unscientific models (like Freud's psychosexual stages) and tests (like the Thematic Apperception Test) (Capara, 1992). After studying World War II veterans in hospitals, Eysenck isolated the variables of extraversion and neuroticism as being key in comprising personalities (Hall & Lindzey, 1985). No matter the person, he or she would fall

somewhere along a spectrum of extraversion-introversion and normality-neuroticism. After more exploration through factor analysis, Eysenck proposed a three-factor model of personality that included extraversion, neuroticism, and psychoticism (Boyle et al., 2016). According to Eysenck, these personality factors are dimensional in nature, with every individual falling somewhere along their continuum (Netter et al., 2020). Eysenck described extraversion as being lively and responsive while introversion was typically defined as being reliable and thoughtful; neuroticism was defined as having below-average emotional control; and psychoticism was labeled as insensitivity, cruelty, a disregard for danger, and a liking for unusual things (Goldberg & Rosolack, 1994).

As each factor is based on a continuum, an individual could be high in neuroticism and extraversion, which could manifest itself as being touchy and aggressive, or high in neuroticism and introversion, which could come across as anxious and pessimistic (Hall & Lindzey 1985). Eysenck's theory built upon those posited by Freud and the behaviorists of the day, as he took the variable of neuroticism, which Freud also discussed, acknowledged the environment in a way similar to behaviorists, but framed it in a new way: that neuroticism is a conditioned fear reaction. Clearly, despite the field refining their measurement and trying to be grounded more in only scientifically based evidence, the great debate regarding the role that the environment plays on an individual's choices still had not been settled; each theorist revisited the same topics with slightly different ideas. Like Eysenck's theory, the five-factor model (FFM) of personality was derived from factor analysis and revolutionized the field (Goldberg, 1990). The FFM was first introduced in the 1930s, when Thurstone had individuals rate their peers from a list of 60 adjectives, which he found could be accounted for by five common factors (Goldberg, 1990). Throughout the years, multiple researchers had dipped in and out of the FFM, but none

immersed themselves until the 1980s when Costa and McCrae created the first FFM inventory (Goldberg, 1990). In this model, the whole universe of trait dimensions was reduced to five main bipolar continuums (McAdams, 1997; Perkmen et al., 2018). These factors were named "the Big Five", and consisted of neuroticism, extraversion, openness to experience, conscientiousness, and agreeableness.

In this model, neuroticism and extraversion are defined similarly as in Eysenck's model. However, openness to experience is detailed as the propensity toward open-mindedness, curiosity, sophistication, and creativity (John & Srivastava, 1999). Those who score high in agreeableness are often described as warm-hearted, forgiving and trust-worthy (Costa & McCrae, 1992; Goldberg, 1990). Lastly, people who score high in conscientiousness are usually reliable, disciplined, and detail-oriented individuals (Costa & McCrae, 1992; John & Srivastava, 1999). Eysenck disputed the need for these extra personality variables, claiming that they are not basic dimensions of personality, but instead extra variables or second-order factors (Eysenck, 1992). However, Costa and McCrae responded with "the basic dimensions of personality are those which together summarize with maximal efficiency the covariation among all the traits in the personality sphere; they must account for all the major variables that have been studied by psychologists as well as those traits that are used by laypersons to characterize themselves and their acquaintances. All five factors are needed to do that" (Costa & McCrae, 1992, p. 656). The FFM has been cross validated among samples of different ages, sexes, races, and languages (Qing Zeng et al., 2003; Zheng et al., 2008). Regardless of race, gender, sex, sexual orientation, and culture, FFM theorists believe that each personality will be comprised of the same five factors, only every individual falls somewhere different on each continuum. Heritability estimates as large as 61% suggest that there may be biological basis for these personality traits,

embedded in genetics or in neurophysiology (Costa & McCrae, 1992; Jang et al., 1996; Perkmen et al., 2018). Some studies have shown that the needs Murray listed are correlated with the FFM; for example, conscientiousness has been positively correlated with achievement, and agreeableness has been positively correlated with nurturance (Billstedt et al., 2017). In this way, the personality theorists of old may have shaped some of the thought in the field at present. The Five Factor model has helped the field of personality psychology along its journey toward becoming a more empirically based field: Cronbach even described the FFM as uniting a somewhat chaotic field (Cronbach, 1970). Today, the five-factor model serves as the grand organization upon which other theories are based and is the most widely used model in personality psychology (Dumont, 2010; McAdams, 1997; Soto et al., 2015).

A Closer Look at Neuroticism

Of all the potential personality traits, there is least disagreement in the field about the existence of neuroticism. Neuroticism is included in most personality measures, including both Eysenck's dimensional personality theory and Costa and McCrae's FFM (Costa et al., 1986). In the Neuroticism Extraversion Openness Five Factor Inventory (NEO-FFI; McCrae & Costa, 2004b), neuroticism is on a bipolar continuum where neuroticism is on one end and emotional stability is on the other, suggesting that the opposite of emotional stability is neuroticism (McCrae & Costa, 2004a). Neuroticism includes behavioral components like being impulsive and highly strung, and cognitive components like worrying, feeling jealous, insecure, tense, self-conscious, vulnerable, impulse-ridden, and emotional (McCrae & Costa, 1987; McCrae & Costa, 2004b). These definitions are reflected in the neuroticism questions from the BFI-2, which includes "can be tense", "worries a lot" and "often feels sad" (Soto & John, 2017). Often, neuroticism is described as negative emotionality, or the propensity to experience negative

emotions like anxiety, depression, anger, or embarrassment (Costa et al., 1986; Soto & John, 2017).

Given that neuroticism is synonymous with negative emotions, neuroticism has been marked as a risk factor for depression and anxiety (Yoon et al., 2013). For example, high levels of neuroticism have been shown to predict the onset of depressive episodes, a finding that has been replicated among cross-sectional and longitudinal samples (OR = 1.50, $p<.001$, 95% CI [1.45, 1.56]; Kendler et al., 2006). Further, a meta-analysis found significant effect sizes for the relationship between neuroticism and anxiety disorders (d= 1.04, 95% CI [0.68, 1.61]; Malouff et al., 2005). These findings suggest that those who are high in neuroticism are also more likely to score highly on measures of depression or anxiety. For these reasons, sometimes research into neuroticism includes measures of depression or anxiety to stand-in for pure neuroticism measures (Calmes & Roberts, 2008; Grames et al., 2008; Karney et al., 1994).

Neuroticism, or negative emotionality, gives rise to certain behaviors in everyday life. Distress levels in general are higher for neurotic individuals, so it is typical to see disturbed thoughts and behaviors (McCrae & Costa, 1987). Neuroticism has been linked to inappropriate coping responses like self-blaming and avoidance, and longer periods of intense anger (Atkinson & Violato, 1994). Neurotic individuals tend to adopt irrational beliefs, have poor coping efforts, and engage in the tendencies to overeat, smoke and drink excessively (Costa & McCrae, 1980; Costa & McCrae, 1986; Teasdale & Rachman,1983). Using structural equation modeling, neuroticism was strongly associated with the latent variables of anxiety sensitivity (factor loading = .70), intolerance of uncertainty (factor loading = .78), perfectionism (factor loading = .68), and experiential avoidance (factor loading = .63) (Naragon-Gainey, & Watson, 2018). Other correlates of neuroticism include being female (r = .20), having higher rates of job burnout

(r^2 = 53.46%), lower levels of job satisfaction (r = -.13), greater difficulty quitting smoking, writing longer posts on social media (r = .10), and an overall different learning style than their non-neurotic peers (Bianchi, 2018; Byrom & Murphy, 2013; Shen et al., 2015; Törnroos et al., 2019; Zvolensky et al., 2020). To summarize, individuals high in neuroticism have greater difficulty feeling happy, accepted, and secure.

How is Neuroticism Measured?

Neuroticism is typically measured through one of two ways: a general personality test, or a test of one or some of the traits subsumed by neuroticism like depression or anxiety (as depression and anxiety are considered components of neuroticism). Some of the general personality inventories include the Big Five Inventory (BFI; John et al., 1991), the Revised NEO Personality Inventory-Revised (NEO-PI-R; Costa & McCrae, 1992), the NEO-FFI (McCrae & Costa, 2004b), or the Eysenck Personality Questionnaire-Revised (EPQ; Eysenck & Eysenck, 1993). Each of these tests measures the various personality factors with each associated theory- Costa and McCrae's "Big Five" theory, or Eysenck's 3-factor theory- and does well psychometrically (Magalhães et al., 2014; Scollon & Deiner, 2006).

Rooted in the Big Five personality theory, the BFI-2, NEO-FFI, and NEO-PI-R all have items related to extroversion, neuroticism, agreeableness, openness to experience and conscientiousness (Costa & McCrae, 2017; Soto & John, 2017). Items are typically high in face validity. As such, it is common to see questions like "has a soft heart" for the agreeableness subscale, or "is outgoing and sociable" on the extroversion subscale (Soto & John, 2017). The BFI-2 and the NEO-FFI both have 60 items which are answered on a 5-point Likert scale (McCrae & Costa, 2004b; Soto & John, 2017). Much greater in length, the NEO-PI-R has 240 items that are answered on a 5-point Likert scale (Costa & McCrae, 1992). Respondents aged 14

and above are considered appropriate test takers for all three personality tests (McCrae & Costa, 2004b). As could be expected, the BFI-2, NEO-FFI and NEO-PI-R share high correlations, demonstrating good convergent validity with one another (Costa & McCrae, 2004; Costa & McCrae, 1992; Rammstedt et al., 2020).

On the other hand, the EPQ is rooted in Eysenck's 3-factor theory and has 100 items (Eysenck & Eysenck, 1993). The EPQ is comprised of an extraversion subscale, a psychoticism subscale, a neuroticism subscale, and a lie subscale (Eysenck & Barrett, 2013). Eysenck defines extraversion as someone who is sociable, craves excitement, and likes to laugh or play practical jokes (Almiro et al., 2016; Eysenck & Barrett, 2013). The psychoticism subscale includes items related to being cold, hostile, and having a liking for odd and unusual things (Almiro et al., 2016; Eysenck & Barrett, 2013). The neuroticism subscale measures items related to being worried, moody, and highly strung (Almiro et al., 2016; Eysenck & Barrett, 2013). Lastly, the lie scale measures if social desirability has clouded an individual's answers (Almiro et al., 2016; Eysenck & Barrett, 2013). Answers are dichotomous (yes/no) and higher scores indicate higher presence of socially desirable responding (Vazquez et al., 2019). For scenarios where it is unfeasible to use the full-length original edition, there is a short-form of the EPQ, which is a 24-item self-report questionnaire that still has the same four subscales (Vazquez et al., 2019).

As far as the advantages and disadvantages of each test, the BFI boasts short, simple items that avoid complex sentence structures (John et al., 1991), Each scale has a similar internal consistency score: the BFI has a Cronbach's alpha of .84, the EQP has shown Cronbach's alpha ranging from .75-.81 (Karney & Bradbury, 1997; Lavner & Bradbury, 2012; Scollon & Diener, 2006; Vázquez et al., 2019), while the NEO has a Cronbach's alpha of .81 (Magalhães et al., 2014). Concerning validity, the BFI and the NEO-FFI have shown high convergent validity with

the NEO-PI-R as scores show correlations as high as .88, which means that test takers scored similarly on both tests (John et al., 1991; Magalhães et al., 2014). The EPQ has demonstrated construct validity through confirmatory and exploratory factor analyses that consistently demonstrate a four-factor solution (Almiro et al., 2016). Additionally, high correlations between the EPQ's 3 clinical subscales and corresponding MMPI items have provided more evidence for convergent validity ($r = 0.8$, $p<.001$ for the neuroticism subscale; Gentry et al., 1985). Regarding cross-validation, the EPQ has shown evidence for reliability and validity across samples from 35 counties, including the USA, Australia, China, Israel, Northern Ireland, and many more (Almiro et al., 2016; Eysenck & Barrett, 2013). To summarize, each of the leading personality tests does very well psychometrically.

As discussed above, neuroticism gives rise to feelings of depression and anxiety. Even the BFI-2 neuroticism domain scale has an anxiety subscale and a depression subscale (Soto & John, 2017). Thus, instead of using a personality measure, it is common to see neuroticism research use depression or anxiety measures. Widely accepted depression or anxiety tests that are most commonly seen in neuroticism research include the Beck Depression Inventory (BDI; Beck et al., 1961), the Center for Epidemiology Studies-Depression scale (CES-D; Radloff, 1977), and the Beck Anxiety Inventory (BAI; Beck et al., 1988). The BDI has excellent internal consistency (α=.93) and has demonstrated construct validity through moderate correlations with scores on other measures of internalizing symptoms, and low correlations with scores on other measures of externalizing symptoms (Beck et al., 1996). Similarly, the CES-D has a Cronbach's alpha of .89, and validity has been reported through confirmatory and exploratory factor analyses (Boulard et al., 2014). For the BAI, researchers have reported a Cronbach's alpha of .92, good test-retest reliability, and a four-factor solution in confirmatory factor analysis, suggesting strong construct

validity (Calmes & Roberts, 2008; Nordhagen, 2001). Due to the substantial evidence that depression and anxiety are components of neuroticism, it is believed that they will measure neuroticism in this study with as much precision as a "pure" neuroticism measure.

Relationship Quality

As discussed above, neuroticism and relationship quality are strongly related to one another, but the exact reasons why are still being debated (Karney & Bradbury. 1997). The second variable in this meta-analysis, relationship quality, can be explained by how good a relationship is, with all things considered. Communication, strain, support, and stability have all been labelled as essential components of a relationship that describe the overall quality of the relationship (Roberson & Fincham, 2018). In the Quality Marriage Index, a measure of marital quality, multiple questions are aimed at commitment and satisfaction, i.e. "Our relationship is very stable" and "This relationship makes me happy" (Norton, 1983). Thus, relationship satisfaction and relationship commitment play a role in relationship quality (Givertz et al., 2009). They are not the same, yet they all exert some degree of influence on the other and play a role in an individual's decision to stay or leave a relationship.

Relationship commitment and relationship satisfaction are variables that are similar to relationship quality—and some evidence even suggests that they are intertwined—but are slightly different in their definitions. First, relationship commitment is the act of pledging or engaging oneself to a partner and can be obligated by pledge, like marriage, or assurance (Pryor & Roberts, 2005). Often, the level of relationship satisfaction can determine the level of relationship commitment, but not always (Givertz et al., 2009). Conclusions from multiple studies have delineated three dimensions of relationship commitment: 1) personal commitment, which arises from the positive aspects of the relationship and leads a person to simply desire to

stay in a relationship; 2) moral commitment, which arises from a partner's religious or moral beliefs about separation and the sanctity of the relationship; 3) structural commitment, which arises from a partner's perceived costs and benefits of staying in the relationship (Adams & Jones, 1997; Johnson, 1991). For all these reasons, an individual may choose to stay committed to the relationship or may choose to break their commitment and leave the relationship.

In a similar vein, relationship satisfaction is defined as the degree to which an individual feels happy within his or her relationship (Moini, 2017). Relationship satisfaction can be impacted by factors outside of the relationship – namely, attachment style, differentiation (or the ability to emotionally and physically separate oneself from one's family of origin), and personality (Alexandrov et al., 2005; Moini, 2017; Miller et al., 2004; Skowron, 2000). Securely attached individuals have higher levels of relationship satisfaction, whereas insecure avoidant or anxious individuals are more prone to jealousy, sensitivity, and withdrawal (Branchaud, 2019; Meyers & Landsberger, 2002). Ultimately, securely attached individuals have a lower divorce rate (Branchaud, 2019). Differentiation, or the ability to emotionally and physically separate oneself from one's family of origin, has been shown to be associated with higher levels of relationship satisfaction (Skowron, 2000). Lastly, personality style is correlated with relationship satisfaction. Those who are high in agreeableness and conscientiousness and low in neuroticism tend to have the highest levels of relationship satisfaction (Razeghi et al., 2011).

Some studies have explored exactly how these three constructs are related to one another (Givertz, et al., 2009; Zhong et al., 2016). Zhong et al. (2016) found that relationship quality influences relationship stability via relationship commitment (b = .11 for men; b = .09 for wives). To expound, if a partner feels that the quality of their relationship is lower than he or she would like, but still has high levels of commitment to the relationship (perhaps because of moral

or structural commitment), the relationship will typically remain intact (Zhong et al., 2016). The interdependence theory may hold some explanatory power for this relationship: individuals tend to act in a way that will maximize rewards and minimize cost (Kelley & Thibault, 1978). With that in mind, commitment to a relationship may be a function of the levels of relationship satisfaction and also the potential costs that may follow exiting the relationship (Adams & Jones, 1997; Givertz et al., 2009). In light of these findings, it is worthwhile to also measure relationship commitment and relationship satisfaction as well as relationship quality, as both commitment and satisfaction are separate but related to relationship quality.

Relationship quality is typically measured through a few commonly used tests: the seven item Relationship Assessment Scale (RAS; Hendrick, 1988), the six item Quality Marriage Index (QMI; Norton, 1983), the 15-item Marital Adjustment Test (MAT; Locke & Wallace, 1959), the 32-item Dyadic Adjustment Scale (DAS; Spanier, 1976), and the 32-item Couples Satisfaction Index (CSI; Funk & Rogge, 2007). With over 2000 citations, the DAS is the most commonly used test to assess relationship quality; the MAT is second most widely used with 1400 citations; and lastly the QMI is the third most popular with over 200 citations (Funke & Rogge, 2007; Locke & Wallace, 1959; Norton, 1983; Spanier, 1976). Despite its popularity, the DAS has fallen under fire recently for having extremely heterogeneous item content that introduces possible confounding variance from factors like communication skills (Funk & Rogge, 2007). The MAT and the DAS were both designed to distinguish between well-adjusted and distressed partners, and the QMI (which has more global items and is more homogenous in its item content) was written to address the shortcomings in both the MAT and the DAS (Locke & Wallace, 1959; Spanier, 1976; Funk & Rogge, 2007). Interestingly, the CSI was developed through performing item-response theory on existing measures of couple satisfaction (Funke & Rogge, 2007). Item-

response theory is a statistical analysis that analyses item characteristic curves to show the quality and precision of measurement tools (Kean & Reilly, 2014). Powerful and precise, the CSI provides great amounts of information from a small amount of items; even the information gleaned from the 4-item CSI surpasses the levels of information gained by the MAT and the DAS, which are both markedly longer (Funke & Rogge, 2007). Overall, these tests measure the quality of, commitment to, and satisfaction within a relationship.

Psychometrically, these tests each demonstrate good evidence for both reliability and validity (Aníbal González-Rivera, 2020; Freeston & Pléchaty, 1997; Funk & Rogge, 2007; Ripley et al., 2018; Zimmermann et al., 2019). In terms of internal consistency, each test demonstrates excellent Cronbach's alpha (α's ranging from .98 - .88) and extremely strong positive correlations with one another (*r*'s ranging from .78 - .92), suggesting excellent convergent validity.

How Does Neuroticism Affect Relationship Quality?

As discussed above, neuroticism is the trait that is most strongly associated with lower levels of relationship quality (Karney & Bradbury, 1997). Only two meta-analyses have previously studied each of the "Big Five" personality factors and how they relate to relationship quality. The most recent of the two found a negative relationship between neuroticism and relationship quality (r = -.44, 95% CI [-.27, -.60]) from a total of 18 studies conducted only in Iran (Sayehmiri et al., 2020). A second meta-analysis found a similar but smaller finding (r = -.22, 95% CI [-.26, -.19]) among 10 studies examining heterosexual relationships (Malouff et al., 2010). Moderation analyses found that there were no differences in the relationship between neuroticism and relationship quality across measures (i.e., NEO vs. other) (Malouff et al., 2010). Further, cross-sectional vs longitudinal data also were not significant moderators of the

relationship, suggesting that this relationship holds over time (Malouff et al., 2010). However, with only a small number of included studies and little variation in the sexual orientation of couples and geographic location, it is unknown whether these effects generalize to a larger and more diverse sample of studies.

In single studies, research has demonstrated the same negative association between neuroticism and relationship quality regardless of culture, age, sexual preference, and length of relationship (Cundiff et al., 2012; Whitton & Kuryluk, 2014; O'Rourke, 2005; Sharma & Raju, 2013). Samples of couples in India (r = -.28, p < .01), Switzerland (r = -.21, p < .05), Belgium (r = -.18, p <.05*)*, and China (r = -.22, p < .01) each exhibited lower levels of relationship quality when there were higher levels of neuroticism (Decuyper et al., 2012; Miller et al., 2013; Sharma & Raju, 2013; Weidmann et al., 2017). Neuroticism is negatively associated with relationship quality in samples of college students aged 19 (r = -.31, p < .001), married couples aged 41(r = -.26, p < .01), and married couples aged 62 (r = -.24, p < .001; Beach et al., 2003; Braithwaite et al., 2016; Cundiff et al., 2012). The same negative relationship exists for homosexual couples as well as heterosexual couples (r = -.31, p < .01; Whitton & Kuryluk, 2014), and whether the sample had been in a relationship for 19 months (r = -.24, p < .01), 23 years (r = -.21, p < .05), or 38 years (r = -.31; O'Rourke, 2005; Robins et al., 2002; Weidmann et al., 2017). Another meta-analysis is needed to examine additional studies that illuminate the role of culture, age, sexual preference and length of relationship.

Both intrapersonal and interpersonal models of neuroticism offer explanations for how neuroticism might decrease relationship quality. Intrapersonal models of neuroticism are focused on the behaviors or personality traits that individuals bring into the relationship with them (Whitton & Kuryluk, 2012). According to the intrapersonal model, neuroticism colors the way an

individual perceives his or her life; individuals who are highly neurotic have a trait-like tendency "to report distress, discomfort, and dissatisfaction over time regardless of the situation" (Karney & Bradbury, 1997; Watson & Clark, 1984, p. 483). So, those who are high in neuroticism are more likely to perceive their lives negatively, become less satisfied with their lives and in turn, less satisfied with their relationships (Cote´ & Moskowitz, 1998). Very few studies have investigated how the intrapersonal aspects of neuroticism affect relationship quality (Whitton & Kuryluk, 2012).

By contrast, the role that interpersonal factors play in the association between neuroticism and relationship quality has been studied to a larger extent. Interpersonal models of neuroticism and relationship quality are rooted in the interactions that take place between partners (Karney & Bradbury, 1997; Schaffhuser et al., 2014). This model claims that spouses learn whether or not they are in a satisfying relationship based on the interactions they have with each other on a daily basis (Gottman, 1990). The judgments each partner makes then influence subsequent behaviors within the relationship – thus, interactions are the mediator between neuroticism and relationship quality (Bradbury & Fincham, 1991). Interpersonal models of neuroticism also suggest that those high in neuroticism create their own negative life experiences through their behaviors and their emotions (Karney & Bradbury, 1997). In a relationship, this could appear as one neurotic partner having an overly emotional response to a neutral event, leading to an argument between the partners which decreases the relationship satisfaction. Observational research supports the notion that interactions affect behavior which affects relationship quality over time (Gottman, 1990; Karney & Bradbury, 1995).

Likely, elements of both the interpersonal and intrapersonal models lead to lower levels of relationship quality. Similar to personality theories of old, dynamic interactionism states that

there is an ongoing transaction between the person and the environment, where an individual's personality shapes the experience, interpretation and reaction to the environment, (Schaffhuser et al., 2014). In this, the individual's personality characteristics will shape their environment, and the environment will shape the individual's personality characteristics. Social exchange theory, which states that individuals evaluate their commitment to a relationship based on the costs and benefits of that relationship, is an example of the intrapersonal model with interpersonal aspects: an individual's perceptions of the costs and benefits of their relationship is extremely intrapersonal, yet the couples' interactions are interpersonal in nature (Fisher & McNulty, 2008). In these ways, it can be extremely difficult to separate the two models of neuroticism for research purposes and posit that only one is the most valid, as likely they are intertwined.

Based on the interpersonal theory of neuroticism and relationship quality, one's partner's neuroticism can have just as detrimental an effect on the relationship as one's own neuroticism (Karney & Bradbury, 1997). Much research into neuroticism and relationship quality classifies neuroticism as either actor neuroticism or partner neuroticism. The effect that an individual's own neuroticism has on his or her relationship satisfaction is referred to as the "actor effect", while the effect that the partner's neuroticism has on the individual's relationship satisfaction is referred to as the "partner effect". For example, if a woman is highly neurotic and this decreases her own relationship satisfaction that is an actor effect. However, if a woman is highly neurotic and this decreases the man's relationship satisfaction as a result, that is a partner effect.

Partner effects of neuroticism have been shown to have just as strong an effect on relationship satisfaction as actor effects (*r*s range from -.19 to -.41; Karney & Bradbury, 1997; Kurdek, 1997; Robins et al., 2002; Russell & Wells, 1994b). Shoda et al., (2002) proposed a cognitive-affective behavior system which states that behavior is determined largely by the

characteristics of the individual and the environment. Within a partner effect, the partner's neuroticism can be classified as part of the environment, which then interacts with the other partner's personality and can lead to poorer interpersonal functioning (Fisher & McNulty, 2008). In these ways, neuroticism can affect both the individual experiencing the neuroticism and his or her partner.

A Case for Causation

Despite repeatedly observing a strong negative correlation between neuroticism and relationship quality in the literature, we must be mindful that "correlation does not equal causation". While it is true that correlation alone does not equal causation, correlation in conjunction with other criteria can make a strong argument for causation. In 1965, Bradford Hill outlined an influential framework with principles and criteria that need to be met in order to establish causation (Hill, 1965). A strong association between variables is the first criteria that needs to be met: if two variables are not associated, one likely does not cause the other. Second, this association must be consistently found among various populations, at different time points, and from different research labs. Third, one variable must precede the other, as cause requires variable A to occur in time before variable B. Fourth, a greater dose of variable A must produce a stronger effect on variable B, for example, higher levels of neuroticism must produce lower levels of relationship satisfaction. Fifth, an experimental design with variable manipulation is needed in order to know that A causes B, although this isn't always possible to do. Sixth, a specific cause must produce a specific effect. Seventh and last, there must be sufficient plausibility to the theory; the theory must be logical and make sense in light of all other things that are currently known about the particular construct.

As the above literature review suggests, neuroticism and relationship quality have been found to be strongly negatively associated with one another, and this association has been found at different time points, among different populations and from different research labs (Braithwaite et al., 2016; Decuyper et al., 2012; O'Rourke, 2005; Whitton & Kuryluk, 2014) . However, as yet, there are no reviews that look specifically into the existing studies on neuroticism and relationship quality to see if the criteria for causation has been met. It is true that some facets of the criteria might be out of the question due to ethical considerations. For example, participants cannot be subject to experimental manipulation to make some higher in neuroticism and others, nor can researchers assign participants to high relationship quality and low relationship quality conditions. However, the following meta-analysis will take place with these criteria in mind to explore which criteria have already been met and provide future direction for the field to satisfy feasible criteria in future studies.

The Current Study

As explained above, many studies have already explored and replicated the correlational relationship between neuroticism and relationship quality, which is the first two of Hill's criteria for causation (Beach et al., 2003; Chen et al., 2007; Karney & Bradbury, 1997). That said, when taken alone, these studies are not immune to the limitations so commonly seen in research: a small sample size, measurement that only reports internal consistency reliability, and lack of ability to infer a causal relationship. Card (2015) made the case that meta-analysis is a powerful tool to draw conclusions and overcome obstacles in the current literature. I can better explore and understand the typical findings and sources of variability across studies by systematically reviewing results from numerous studies (Card, 2015). The purpose of this meta-analysis will be to address the above limitations by quantifying the correlational relationship between

neuroticism and relationship quality using studies from across diverse cultural and ethnic backgrounds, different measurement tools, and some longitudinal data. In our analyses, I will take into account the cultural background, length of relationship, age of participants, measurement tools used, experimental design, and whether the data are cross-sectional or longitudinal to discuss current findings in light of Hill's criteria for causation. This is to detail whether neuroticism and relationship quality have been studied enough to infer causation. Overall, I hypothesize that there will be a negative relationship between neuroticism and relationship quality regardless of these variables. By looking at the field as a whole, it is our hope to summarize and synthesize the existing literature to develop a model about how neuroticism operates in relationships, comment on causation, and chart a course for necessary future research and measurement.

Method

Data Storage and Management

I obtained relevant articles through two approaches. First, I typed variations of the following search terms into PsycINFO, PsycARTICLES, Academic Search Premier and GoogleScholar: "Neuroticism AND Marital Satisfaction", "Neuroticism AND Marital Quality", "Neuroticism AND Relationship Satisfaction", "Neuroticism AND Relationship Quality", "Neuroticism AND Satisfaction", "Personality AND Relationship Satisfaction", Personality AND Relationship Quality", "Personality AND Marital Satisfaction", "Personality AND Marital Quality", and "Personality AND Satisfaction." There was no limit on publication year to ensure the inclusion of every study performed on these topics. Second, I consulted the references of retrieved articles to find earlier relevant studies. Once a list of titles had been compiled, each abstract was reviewed by a group of 6 undergraduate research assistants who checked for

measures of neuroticism (including depression or anxiety) and relationship quality (including satisfaction and commitment). If both constructs were measured and the article is written in English, the article was saved in a database for coding.

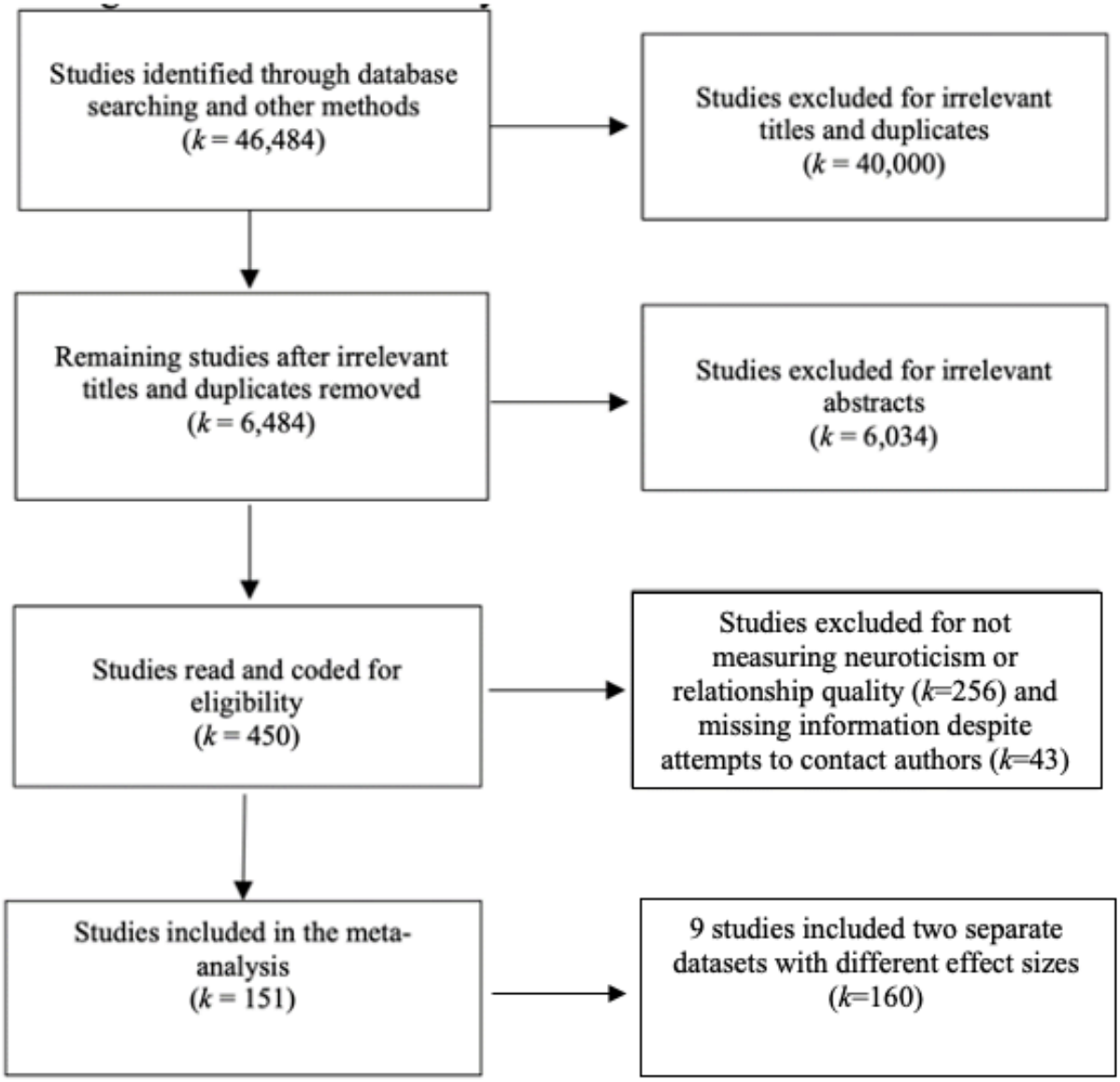
Our initial search yielded 46,484 results. I screened these and excluded 40,000 titles that were irrelevant to the topic or were duplicates. The remaining 6,484 were screened for relevant information in the abstract, yielding 450 full-text articles which I assessed for eligibility for inclusion. It is possible that I missed some studies that reported information on neuroticism and relationship quality but did not include that information in the title or abstract. 256 of those articles were excluded because they did not fit the inclusion criteria, leaving 194 articles. Of those 194, I had to email 68 authors for missing data and further excluded 40 studies due to lack of reply. Further, 3 authors responded that they did not have the missing data, so those studies were excluded. I included the remaining 151 studies in our analysis.

Inclusion and Exclusion Criteria

An article was included in the meta-analysis if it was accessible in English, and if it contained a measure of neuroticism, depression, or anxiety, and relationship quality, relationship satisfaction, or relationship commitment. Last, I only included studies that used samples who were in a romantic relationship: I did not include any studies that reported a measure of neuroticism and relationship quality among friendships or familial relationships. If an article contained a measure of both variables but did not report a correlation or regression coefficient of those variables, I emailed the authors to obtain that effect size. If they did not respond, I excluded the study due to not having the necessary information for inclusion. Publication was not a necessary requirement for inclusion.

Figure 1

Flowchart for Inclusion/Exclusion Decisions



Coding

Articles included in the meta-analysis were coded twice by undergraduate research assistants who received individual and group training meetings each month to increase the

reliability of their ratings. Prior to beginning solo coding work, I provided a thorough review of the coding process, including what variables to look for and how to locate hard-to-find statistics. Each undergraduate research assistant coded two articles and was given specific feedback before being allowed to work on the rest of the articles. From then on, I performed monthly checks of all articles coded to ensure reliability. Articles were coded using a google form that asked questions like "What is the title of the study" and had a line to submit a response (see Appendix A). Each response went on a google sheet which became our initial database. Once each article had been coded twice by different research assistants, I went through each submission on the google database to check that both articles were coded to contain the same answers. In addressing any discrepancies, I referred to the original source to find the answer, or I contacted the authors if the manuscript did not contain the answer.

Moderators and Study Characteristics

I extracted the following information from the methods and results sections of each study: year of publication, publication source, how the study was funded (if the funding was listed, I copied and pasted the source and/or grant number; if funding was not listed, I wrote "N/A"), the sample size for men, sample size for wives, the country that the sample was from, mean age for men, mean age for women, sexual and racial minority percentages, average combined family income, average years of education for men, average years of education for women, the percentage of couples that had children, and average length of relationship in months. For the purposes of this study, our sexual minority variable was defined as the percentage of participants who identified as gay, lesbian, bisexual, or asexual. If the manuscript did not include that information, I coded it as 0%. Similarly, our racial minority variable was defined as the percentage of participants who identified as Hispanic, Latino/Latina, Asian,

Polynesian, or African American. If a manuscript did not include this information, I coded racial minority as 0%.

Lastly, I coded methodological factors such as whether the data was self-reported or partner-reported, which measures the study used, and their reliability evidence. For reliability, I coded Cronbach's alpha or test re-test reliability coefficients. If these were not included, I coded the reliability variable as "Not Reported" in our database. I coded whether the study was cross-sectional or longitudinal - if the study was longitudinal, I noted the time-lag in months between the assessments, how many participants remained throughout the length of the study, and what method the researchers used to address data missingness (i.e., full information maximum likelihood or list-wise deletion). I included male and female neuroticism and relationship quality correlations at both the actor and partner levels. If a study did not include a correlation for these variables but instead reported a different metric, I first contacted the study author to ask for the correlation. However, if the author did not respond or did not have a correlation, I followed Peterson and Brown's approach to transform the coefficient provided in the paper into Pearson's *r* (2005). Peterson and Brown found that transforming different metrics into the desired metric rather than simply excluding a study is more likely to reduce bias in the meta-analyzed results (2005). All in all, there were only 12 studies that had to be transformed into Pearson's *r* using this approach.

Finally, I also coded each study for some of Hill's criteria for causation. To do this, I included four columns at the end of our database, and each study received a "1" in the column if 1) there was a correlation between the two variables (indicating strength of association); 2) the data are from different countries, different research labs, or different decades in time (indicating consistency); 3) one variable precedes another variable in any of the studies (indicating

temporality); 4) the study used an experimental design rather than an observational design (indicating experimental evidence). The comparison study I used for criterion two was Karney and Bradbury's 1997 article entitled "Neuroticism, marital interaction, and the trajectory of marital satisfaction", so if a study was from outside of the USA, or had authors from a different research lab, or was from a different decade than 1990–1999, it would receive a 1 in its column.

At the conclusion of coding, I summed each column to see how many studies in the literature meet the requirement for that particular component of causation, to say how many of the criteria for causation have been met. Due to its subjective nature, I was unable to code some of the criteria for causation, including plausibility, whether greater levels of neuroticism results in lower levels of relationship quality, and whether specific cause equals specific effect. Despite not being able to feasibly code for each of the criteria for causation, I feel it is a step in the right direction toward being able to comment on the next direction for the field as far as establishing causality.

Eliminating Bias

To eliminate the bias of the file-drawer effect, I contacted 22 authors with multiple publications in this field to see if they had any currently unpublished studies. All authors reported that they did not have any unpublished data in this area, so no extra studies were added to the database for coding. Due to this, every study included in this meta-analysis is published. In order to test for publication bias, I computed and reported Duval and Tweedie's *Trim and Fill* (2000) method which estimates an unbiased effect size by trimming the most extreme studies of the funnel plot and re-computing the effect size until the funnel plot is symmetrical.

Power

To ensure sufficient power to detect small effects, I performed an a priori power analysis using a simple formula found in Borenstein et al., (2009, pp. 270–272). From looking through studies I had already gathered for this meta-analysis at that point, I estimated having around 106 included studies with an effect size of 0.2 and an average sample size of 834. Under these conditions, power was determined to be 100%.

Data Analysis

To analyze data, I used Comprehensive Meta-Analysis Version 3 to report effect sizes for neuroticism and relationship quality, moderator analyses, and the heterogeneity of included studies. Effect sizes were first transformed into Fisher's z scores, then each analysis was conducted using the Fisher's z scores. Prior to reporting the findings, the scores were transformed back into Pearson's *r* correlations. In the following section, I will report the overall relationship between neuroticism and relationship quality, and separate Pearson's *r* correlations for the man actor correlation, woman actor correlation, man partner correlation, and woman partner correlation. Due to the vast number of studies looking into neuroticism and relationship quality, I used a random-effects model to aggregate study-level correlations, as likely, the true effect size of neuroticism and relationship quality will vary across studies (Borenstein, 2019). This model is especially desirable over a fixed-effects model for applied, real-world data like relationship quality (Borenstein, 2019).

For moderation analyses, I used two-way tests to test for differences between samples from different countries (USA vs non-USA Eastern and non-USA Western), measures (personality vs depression/anxiety), and whether the same relationship exists in longitudinal data and cross-sectional data. Then, I used meta-regression to test for moderation between continuous variables, like whether length of marriage or age of participant changes the association between

neuroticism and relationship quality. To use these variables as moderators, I made sure that there were at least five studies that reported the necessary data (e.g., five studies that report cross-sectional data and five studies that report longitudinal data). To quantify between-study heterogeneity, I used the I^2 statistic, which indicates the percentage of variance attributable to study heterogeneity rather than chance, or a ratio of true to total variance (Borenstein et al., 2009). Or, in other words, I^2 describes the percentage of variation across studies that is attributable to an actual difference in the study's findings rather than chance, on a scale of 0-100. I^2 statistic is a proportion, not an absolute value, that displays the dispersion of true effects, or how much the observed values vary from the true value. An I^2 statistic that is high represents very little variation from the true population value, represented by a narrow, tight distribution (Borenstein, 2019).

Results

Overall Effect Sizes

First, I report on the general effect sizes between neuroticism and relationship quality, then specific to male-actor, male-partner, female-actor, and female-partner. The overall effect size between relationship quality and neuroticism was r = -.222 (k = 160; 95% CI = [-.249, -.195]; p < .01; I^2 = 97.068), suggesting that the aggregate effect size taken from 160 independent samples shows that as neuroticism increases, relationship quality decreases by .222. An I^2 of 97.068, means that sampling error explains 3% of the observed variance, while the other 97% reflects variance in true effects.

The effect size between relationship quality and neuroticism for just-male actor correlations was r = -.225 (k = 94; 95% CI = [-.250, -.199]; p < .01; I^2 = 82.428); for just female-actor correlations was r = -.223 (k = 96; 95% CI = [-.267, -.178]; p < .01; I^2 = 95.217); for just

male-partner correlations was r = -.161 (k = 61; 95% CI = [-.189, -.133]; p < .01; I^2 = 81.169); and lastly for just female-partner correlations was r = -.189 (k = 57; 95% CI = [-.217, -.160]; p < .01; I^2 = 80.972). Taken together, these results show that across a large sample of studies, a negative relationship between the two variables exists at around the .20 level, regardless of whether I look at an overall relationship or whether I look at simply a man, or a woman, at the actor or partner level.

The uniformly large I^2 values (>.50) indicate a large amount of systematic heterogeneity, which make them appropriate for moderator analyses. As I^2 indicates the proportion of true variance that is due to systematic variance rather than sampling variance (a source of error), where there are high I^2 values, I can use moderation analyses to try to discover the sources of the systematic variation, or true effects.

To check for bias in my effect sizes, I ran a Duval and Tweedie trim and fill analysis. This method suggested that 0 studies were missing and the imputed point estimate was -0.222 (95% CI = [-.248, -.195]). From this, I can assume that the addition of any missing studies is unlikely to change the overall effect size. Second, I ran a "leave one out" analysis, which is a sensitivity analysis to see if any one study has an undue impact on the overall effect size. This analysis showed that no study had an undue influence on the reported effect sizes. See Appendix B for a table containing each study and its associated effect size.

Moderator Analyses

In this section, I report the findings from planned moderator analyses of specified variables that could account for some of the substantial systematic heterogeneity observed in the

meta-analyzed effect size. I performed these analyses on the overall effect size (r = -.222) since there was little difference between the overall effect size and the subgroup effect sizes (woman actor, man actor, etc.). Using the overall effect size offers greater power than analyzing the subsets, which did not differ markedly from the overall effect size.

Cross-sectional vs Longitudinal

I compared the effect sizes from cross-sectional correlations to the effect sizes from longitudinal correlations. Results showed a significant difference, however, correlations for both groups were similar to the above overall correlations (r_{cross} = -.199, p = .000, k = 110; r_{long} = -.270, p = 0.000, k = 48; Q = 11.95, p = 0.003). These findings suggest that while there is a statistically significant difference between cross-sectional and longitudinal correlations regarding the association between neuroticism and relationship quality, the difference is small. That is, the effect size obtained from longitudinal data gives a little more weight to the correlation. Interestingly, the effect size from cross-sectional studies trends downward when compared to our overall aggregated effect size, while the effect size from longitudinal studies trends upward, suggesting that neuroticism could have more of an effect on a relationship over time. This could mean that neuroticism unfolds over time and creates a chronic strain on the relationship that cross-sectional findings cannot capture. However, the negative association does not change regardless of whether the correlation is cross-sectional or longitudinal.

Sample Country of Origin

I compared the effect sizes from samples where the country of origin was the United States of America to samples labelled as “Non-USA Eastern” (which included samples from the

continents of Asia and Africa, typically seen as more collectivist cultures where the tradition of arranged marriage may still be commonplace) and "Non-USA Western" (which included samples from the continents of Europe, North America, and Oceania, typically seen as individualistic cultures where the tradition of arranged marriage is not common), resulting in 3 groups for comparison. Results showed a nonsignificant difference between the effect sizes associated with each group (r_{USA} = -.232, p < .01, k = 91; $r_{non\text{-}USA\ Eastern}$ = -.133, p = .097, k = 23; $r_{non\text{-}USA\ Western}$ = -.224, p < .01, k = 40; Q = 3.525, p = .318), suggesting that neuroticism and relationship quality are similarly associated regardless of the country of origin of participants in the sample. The p values associated with the correlations are significant, suggesting a significant relationship between each variable, however the p value associated with the moderation analysis is non-significant, indicating that there is not a significant difference between effect sizes obtained from different countries. From this, I can deduce that country of origin does not moderate the relationship between neuroticism and relationship quality. As I grouped by location based on culture (collectivist vs individualist), we can even say that neuroticism and relationship quality operate similarly among both cultures.

Types of Measurement

I compared the effect sizes of correlations obtained from a personality measure that specifically measured neuroticism to effect sizes of correlations obtained from a proxy measure for neuroticism, like depression or anxiety. Results showed a significant but small difference between the two types of effect sizes ($r_{personality}$ = -.208, p < .01, k = 136; r_{depanx} = -.298, p < .01, k = 20; Q = 12.649, p < .01), suggesting that there is a difference between the effect sizes obtained from personality measures and effect sizes obtained from "stand in" measures of depression and anxiety. One possible explanation for this finding could be that psychopathology measures have

a more pronounced effect because they better capture the effect that depression or anxiety has on relationship quality, rather than simply neuroticism. We have operationalized depression and anxiety as part of neuroticism, as explained above, but perhaps these constructs represent only a very specific part of neuroticism than broad "neuroticism", which measures multiple types of neuroticism like anger, hostility, jealousy, insecurity, anxiety, depression, and others. However, the difference between the correlations is small and demonstrates a similar relationship regardless of whether the construct is depression, anxiety, or broadly neuroticism.

Relationship Length

I employed meta-regression to test the moderating effect of two continuous variables—average relationship length and average age of the sample—on the correlation between neuroticism and relationship quality. I found that for each one unit increase in sample average reported relationship length in years, the correlation between neuroticism and relationship quality slightly but significantly weakened (b = 0.003; 95% CI = [.001, .005]; p = .006). This finding means that with longer time spent in a relationship in a sample, the correlation between neuroticism and relationship quality (r = -.222) slightly decreases, taking the effect size closer to 0 by .003 for each additional year. That is, the longer the average relationship in a sample, the association between neuroticism and marital quality gets slightly smaller (or, closer to zero). These findings illustrate that as time increases, an individual becomes less neurotic and sees subsequent changes in their relationship quality, or that a couple can work through their relationship problems and learn how to navigate one or both partner's neuroticism, increasing their relationship quality.

Participant Age

For this moderation analysis, I found that for each one unit increase in sample average age in years, the correlation between neuroticism and relationship quality slightly weakened (b = 0.001; 95% CI = [-.001, .004]; p = 0.322). However, this relationship was statistically nonsignificant, suggesting that the age of the participant does not reliably moderate the relationship between neuroticism and relationship quality. In other words, neuroticism and relationship quality have the same relationship regardless of the average age of a sample, or that neuroticism doesn't differentially relate to relationship quality based on age. However, an important insight to consider, in light of our above findings, is that the association between neuroticism and relationship quality has more to do with the length of the relationship than the age of the individual. According to the maturation principle, people become less neurotic over time (Marsh et al., 2013), which may counteract a cumulative effect on neuroticism over time. So, simply looking at the age of the individual may not fully capture the same complexities of the relationship between neuroticism and relationship quality as the length of relationship does.

Case for Causation

As delineated in the Method section, I performed additional coding on each study to see if taken together, there is enough literature on neuroticism and relationship satisfaction to be able to make a case for a causal relationship. Of the criteria that I was able to code for, I found substantial evidence that there is a relationship between the two variables, evidenced by a correlation and replicated by this meta-analysis. Each of the 151 included studies reported a correlation between the two variables, suggesting that there is at least an association between the two of them. Second, I coded whether studies from different labs, during different decades, and from different countries (using Karney & Bradbury 1997 as our baseline comparison study) have also found a similar association. I found that 149 of our 151 studies met one of those criteria (a

different lab, a different decade, or a different country of origin), suggesting that these findings hold regardless of who is performing the study, what time the study was held, and where the study was held. Indeed, specific moderator tests of which region of the world the study was conducted provided no evidence for differences.

Next, I coded for temporal precedence of one of the variables by counting how many studies presented longitudinal data and found that 52 of our 151 studies contained longitudinal data. Some studies reported a longitudinal effect size with neuroticism being measured first, and relationship quality being measured later, but others reported effect sizes with relationship quality measured first and neuroticism measured later. So, despite there being evidence for broad temporal precedence, I cannot ultimately say which specific variable precedes the other (neuroticism, or lowered levels of relationship quality).

Last, I coded for experimental design and found that 0 of our studies used an experimental design, for obvious issues as being unable to assign participants to become more neurotic or to disrupt their relationship functioning. Thus, every study we included was observational in design. In order for these criteria to be met, researchers would have to creatively find ways to control more variables than have currently been controlled in past research.

Discussion

The purpose of this study was to systematically aggregate all the research written in English on neuroticism and relationship quality to report overall effect sizes when all studies are combined and weighted between these two variables. I hoped that by using a large sample of studies from diverse backgrounds, I would be able to test the limits of this highly replicated relationship and comment on a possible case for causation. Our findings have replicated the current body of literature by showing another, more precise effect size by nature of the meta-

analytic sample being so large. Further than simple replication, I also hoped to extend the current body of literature by 1) establishing whether criteria for causation has been met, and 2) providing a model for how neuroticism operates through conducting moderation analyses. By exploring both of these extensions of the current literature, I will provide some ideas for future studies in this area.

The results of our meta-analysis confirmed that neuroticism and relationship quality are negatively related to one another. Therefore, as one increases, the other decreases. This could be seen as where there are high levels of neuroticism, there is a subsequent decrease in relationship quality. On the other hand, it could be characterized that as relationship quality increases, neuroticism decreases. The current literature has not been able to parse apart which variable comes first, or if it is more bi-directional in nature. Despite this, the negative relationship between the two variables has been found again and again, replicated in many labs and across many samples, regardless of whether neuroticism is on the actor or partner level, and the country, age, length of relationship of the sample, which measure is used to capture the constructs, and whether data are cross sectional or longitudinal. Even when our moderation analyses showed a statistically significant difference, the nature of the correlation did not change much from the original effect size, suggesting that this relationship can be considered stable.

That said, our moderator analyses contained some findings that add a slight level of nuance to what seems like a robust relationship on the surface. Further exploration of these findings can help us to establish a model for how neuroticism operates within a relationship. First, I found that effect sizes obtained from longitudinal data reported a slightly stronger correlation in the negative direction than effect sizes reported from cross-sectional data. Although both correlations were negative and did not represent a large difference in effect size, their statistically

significant difference tells a story. A possible explanation for this finding could be that neuroticism may not initially corrode a relationship at the outset, but that as relationships continue, the behavior of a neurotic partner may wear down a less neurotic partner and decrease the relationship quality. This particular finding is at odds with Malouff et al. (2010), who found that cross sectional data and longitudinal data were not statistically significant in difference from one another. However, given that Malouff et al. (2010) only used 10 studies in their meta-analysis, our findings might be more accurate due to our large body of studies.

I specifically tested this using length of relationship as a moderator, which demonstrated a statistically significant outcome, suggesting that the length of time someone is in a relationship changes the relationship between neuroticism and relationship quality. However, this finding did not demonstrate the same story as was found with the longitudinal effect size. This result showed that over time, the relationship between neuroticism and relationship quality weakened rather than strengthened. A possible explanation for this finding lies in the maturation effect, which posits that over time, neuroticism decreases, and an individual becomes more agreeable and conscientious (Marsh et al., 2013). Thus, it follows that if neuroticism generally decreases over the life span, the effect size would get closer to zero with each year spent in the relationship. Or, perhaps over time a couple learns to navigate one or both partner's neuroticism, and in turn increases in relationship quality. More pessimistically, it could mean that more neurotic individuals select out of marriage creating a survivorship bias. This finding is not necessarily at odds with our longitudinal moderation effect size, which simply looked at whether the measurement of one or both variables took place in time intervals, most of which spanned only months. It did not capture length of relationship and did not tend to span great periods of time, as did our relationship length variable. One last potential explanation for how longer time spent in

the relationship slowly takes the effect size to zero could be that those who are highly neurotic and experiencing low relationship quality opt out of their relationships quickly, so the individuals making it through longer relationships might naturally be lower in neuroticism and higher in relationship quality.

When I explored the average age of the sample as a potential moderating variable, I found a statistically nonsignificant outcome, which suggests that age of an individual does not change the relationship between neuroticism and relationship quality. This finding offers more evidence for the maturation effect described above, that as an individual ages, their neuroticism likely decreases. Given this information, it is probable that the best way to capture how neuroticism operates over time might be through the length of the relationship rather than simply the age of the individual.

Further, I explored whether measurement type (i.e., a broad personality measure or a measure of anxiety or depression) would moderate the relationship between neuroticism and relationship quality. Different from the findings of Malouff et al. (2010), I found a statistically significant difference, which indicates that the type of measure a study used did alter the relationship between neuroticism and relationship quality. Despite being a very small difference, the measures of depression and anxiety demonstrated a stronger effect size, leaning even more in the negative direction than studies that used a personality measure. This could mean that individuals whose neuroticism leans more in the direction of depression and anxiety rather than other types of neuroticism (like anger, jealousy, hostility etc.) have lowered amounts of relationship quality. When we measure neuroticism broadly, we might get a certain effect size like the ones seen in this study. But perhaps when we start to measure specific types of neuroticism instead of the broad umbrella term of “neuroticism”, we might see slightly varying effect sizes, as this

moderation analysis shows. To fully understand this complexity more, future research could benefit from parsing apart the different types of neuroticism and exploring effect sizes for each type. I likely found this small difference between the types of measurement because of our large number of studies, allowing us significant power to detect small effects. Whereas Malouff et al. (2010) only had 10 studies, so were not as able to perform as thorough of an analysis. Of note, the measures of depression and anxiety may include more psychometrically difficult neuroticism questions, leading to this outcome. Further, these depression and anxiety questionnaires tend to measure state-like current symptomatology rather than broad trait-like neuroticism, so despite being a well-established subcategory of neuroticism, depression and anxiety measures might be getting at a slightly different construct (i.e., acute psychopathology) than broad neuroticism.

Last, I explored whether the sample's country of origin moderated how neuroticism and relationship quality were associated with one another. I found a statistically nonsignificant difference, which means that our two constructs are similarly related regardless of the sample's country of origin. This finding is interesting in light of how I grouped by geographic location and culture. In putting countries together, I tried to group collectivist cultures together and individualist cultures together, largely due to how marriages are valued or entered into in each culture. So, in finding that neuroticism and relationship quality are similarly related despite where a participant lived, I can conclude that this relationship holds across countries and cultures where marriage is viewed differently than in the USA.

Can Causation Be Established?

One of the purposes of this meta-analysis was to provide information regarding whether I have enough data to imply causation between one of these variables. Despite having many, many

studies that investigate neuroticism and relationship quality, I feel that I am as yet unable to make a strong case for causation using Hill's seven criteria for causation (1965). First, as I was unable to objectively code for plausibility, whether greater levels of neuroticism results in lower levels of relationship quality, and whether specific cause equals specific effect, I was unable to thoroughly explore the literature as it relates to all seven of the criteria. However, of the four criteria that I was able to fully measure, our results were still somewhat weak. Since the studies I coded for causation were also included in this meta-analysis, by their very nature they had to include a correlation. So naturally, that criterion was easily met. However, if a study did not report a correlation between the two variables, I was unable to include it in the meta-analysis, which means there is somewhat of a bias in that criterion. Further, using an experimental design would be almost impossible for researchers to employ when studying neuroticism and relationship quality, so we will always be unable to explore this criterion fully.

Establishing temporal precedence was a criterion that I could somewhat explore; but was only obtained through saying whether a study provides data that one variable was measured before the other; this method does not explicitly tell us whether one variable was present before the other, but simply only captures the time of measurement. Last, I was able to fully explore the criterion that states that an effect must hold over time, across labs, and across different countries. I can accurately say that the same association exists between neuroticism and relationship quality across different decades of data, from different researchers, and from samples with different countries of origin. Therefore, I feel that I can only emphatically say that two of the seven criteria for causation have been met (that there is an association between the two variables, and that the same association holds across time, labs, and countries), therefore I hold off on saying

that A) higher neuroticism causes lowered relationship quality, or B) lowered relationship quality causes higher neuroticism.

Future Research

Given the vast number of studies I have gathered for this large meta-analysis, and what I have found in our investigation of a possible case for causation, there are a few directions that the field can take in order to know more about the association between neuroticism and relationship quality. The first avenue for more exploration lies in capturing the temporal precedence of the variables. Currently, I am unable to answer whether heightened levels of neuroticism decreases relationship quality, or whether decreases in relationship quality heighten levels of neuroticism. A study specifically designed to test this could use a cross-lagged panel model to compare participant's scores on relationship quality and on neuroticism, to explore which has a stronger temporal precedence and concurrently, the stability levels of both variables.

Second, simply exploring longer time lags in longitudinal studies is necessary to fully understand the relationship between neuroticism and relationship quality. A majority of the longitudinal studies that look into these two variables only cross a time span of months to perhaps one or two years. To capture the long-term effects of neuroticism, and more fully unpack our finding that length of relationship moderates the relationship, we need more long-term longitudinal data. This could potentially look like collecting population level data on neuroticism levels when participants are teenagers, and relationship quality data in adulthood, including relationship status and relationship quality. Having a widespread sample size coupled with a time lag of years could be very beneficial in unpacking the relationship between neuroticism and relationship quality further.

Another important direction for the field could be to employ as experimental of a design as possible, given ethical constraints. There are multiple ways that the field could inch closer to achieving this goal. One potential way to do this could be, with IRB approval, to assign individuals to perform certain behaviors within their relationships that are linked to neuroticism (e.g., stonewalling, reassurance seeking, starting arguments) while simultaneously measuring relationship quality. However, before employing this type of methodology, the field would need to ascertain the specific behaviors that are linked to higher levels of neuroticism. In order for this to be discovered, we would need to hold a daily diary study where participants who have varying levels of neuroticism keep record of the behaviors they engage in each day, so we can draw correlations between neuroticism and very specific behaviors within a relationship.

Another way the field could inch closer toward a more experimental design could be to randomly assign participants who are high in neuroticism, as shown by evidence-based measures, to perform soothing techniques that will likely lower their emotions and the behaviors associated with their neuroticism. Mindfulness based interventions like meditation, deep breathing, and self-compassion exercises have been shown to decrease the effects of "neurotic anxiety" (Gul & Jahangir, 2019; Stauffer, 2015). With that in mind, the study could measure participants' levels of relationship quality over time while participants perform these mindfulness-based interventions. While this may not meet the definitions of a classic experimental design, it is a step closer to being able to manipulate variables without posing a threat to the ethical commitment of researchers. Evidence for causation is supported when a dependent variable is changed as a result of an independent variable being systematically manipulated, and there are ways to do this without causing harm to a relationship.

Theoretical Underpinnings

The findings of our study fall in line with the assumptions found in the social exchange theory. The social exchange theory posits that couples' evaluate how satisfactory their relationship is based on the perceived rewards and costs of their current relationship. Similar to a business model, partners will try and maximize their transactions for the greatest benefit and least cost (Nakonezny & Denton, 2008). Within the context of the social exchange theory, feeling high levels of relationship quality is the outcome of perceiving greater rewards to staying in the relationship than leaving the relationship; or in other words, the benefits one receives from staying in the relationship are greater than what it is costing the individual to stay in the relationship (Thibaut & Kelley, 1959). "Relationships grow, develop, deteriorate and dissolve as a consequence of an unfolding social-exchange process, which may be conceived as a bartering of rewards and costs both between the partners and between members of the partnership and others" (Huston & Burgess, 1979, p.4). Thus, people approach their relationships the same way they approach other aspects of life - by analyzing the costs and rewards.

The rewards within a relationship could be material in nature like food and protection, or social in nature like companionship, security, frequent sexual intimacy, and support (Bradbury & Karney, 2013). On the other hand, the costs in a relationship could be frequent arguments, physical harm, or financial difficulties. The simple formula "outcome = rewards - costs" can explain why people stay in their relationships. If the rewards outweigh the costs of staying in the relationship, a couple are likely to remain together. However, if the costs outweigh the rewards, they are likely to end their relationship (Thibaut & Kelley, 1959). Simply looking at costs and benefits may not explain the full story, though. Dependence is another factor that can explain why highly unhappy individuals might stay in a relationship (Thibaut & Kelley, 1959). If an individual does not perceive they have many positive alternatives to their relationship, they are

likely to be dependent on it. Furthermore, many relationships have barriers that inhibit exiting, such as facing disapproval from family and friends if the relationship ends in divorce, or fear of potential retaliation from their partner. Individuals also make investments in the relationship, such as a shared home, children, and time spent together that would likely be lost if the relationship ended. Thus, if someone feels that they will not be able to get by without their partner for financial reasons, emotional reasons, or any other constraint they perceive, they will be less likely to leave even an unhappy relationship. Likewise, if they feel that the cost of losing their investments is too great, they are likely to stay in the relationship. These barriers, investments, and lack of alternatives make someone dependent in a relationship, and if a person is highly dependent, they will likely remain in the relationship, even though the costs of staying outweigh the benefits of staying.

Our findings that neuroticism and relationship quality are robustly negatively correlated fall in line with the social exchange theory. Likely, individuals who are highly neurotic are engaging in certain behaviors that are damaging the relationship, or on the other hand, the relationship is suffering and leading people to act in more neurotic ways. Whichever variable comes first, it is at least clear that A) neuroticism harms relationships in some way, or B) poorer relationships lead to higher levels of neuroticism. As neuroticism colors the way an individual perceives their intrapersonal world, and affects their interpersonal style, (Karney & Bradbury, 1997; Whitton & Kuryluk, 2012), it is likely that a highly neurotic individual is either perceiving more costs in their relationship or bringing more costs to their relationship for their partner, or both. In light of the social exchange theory, the way in which the costs outweigh the benefits in the relationship can explain why relationship quality is lower among individuals with high levels of neuroticism. Further, highly neurotic people might feel greater levels of dependence on the

relationship. This could be due to emotional reasons like fear or anxiety about leaving a relationship, or that they perceive they have less attractive alternatives. For these reasons, partners might remain in the relationship despite feeling less satisfied by the relationship.

The social exchange theory, particularly the part about costs and benefits, is consistent with Gottman's *magic ratio* principle. This theory states that there needs to be five times as many positive interactions in the relationship for everyone negative interaction, making the *magic ratio* a 5:1 ratio. That means that for every piece of feedback, criticism, or argument, there needs to be at least five positive interactions "in the bank" like compliments, flirtations, or affectionate touch. Gottman's *magic ratio* is likely off-balanced among partnerships where one or both partners are high in neuroticism, tending toward more negative interactions than positive interactions. As discussed above, highly neurotic individuals may tend to experience more jealousy, more withdrawal, or more arguments in a relationship (Branchaud, 2019; Costa & McCrae. 1986). If a couple is not trying really hard to balance the 5:1 ratio carefully, it would be more likely that there would be greater levels of negative interactions than positive interactions in their relationship. Taken together, both the social exchange theory and Gottman's *magic ratio* hold explanatory power for the negative relationship between neuroticism and relationship quality.

How Does Neuroticism Operate in Relationships?

Given the vast amount of literature I read for this meta-analysis and our findings from moderation analyses, I can infer that individuals who are highly neurotic must be more susceptible to certain thought patterns, behaviors, and ways of interpreting the world around them. As discussed above, neurotic individuals are more prone to worry, jealousy, and insecurity (McCrae & Costa, 2004a). Some individuals display their neuroticism as anger and hostility

(Atkinson & Violato, 1994; Lahey, 2009) while others could display their neuroticism as withdrawal and depression (Soto & John, 2017). For these reasons, I cannot say definitively that neuroticism in a relationship looks like any one single thing.

However, I do assert that there are three avenues through which neurotic individuals experience less positivity and more negativity: through emotions (Yoon et al., 2013, McCrae & Costa, 1987), behaviors (McCrae & Costa, 2004a; McNulty, 2008), and interpretations (McNulty, 2008; Rafienia et al., 2008; Schaffhuser et al., 2014). Specific to a relationship, lowered positive emotions and more negative emotions could look like lower feelings of trust and optimism, and higher levels of depression, anxiety, and anger (Lahey, 2009; McCrae & Costa, 2004a). Further, lowered positive behaviors and more negative behaviors in a relationship could look like less pleasant engagement with one another, constant bids for attention, reassurance seeking behaviors, poor communication, and higher levels of conflict (Iveniuk et al., 2014; Mund et al., 2016; Tong et al., 2018). Last, lowered positive and more negative interpretations could look like a partner interpreting an ambiguous text in a negative light, or not assuming the benefit of the doubt with their partner's intentions (McNulty, 2008; Schaffhuser et al., 2014).

More concrete evidence from future studies is needed to back up these assertions, but based on the large volume of literature I read for this meta-analysis, this is our working model for how neuroticism operates in a relationship. The question as to which comes first—higher levels of neuroticism or lowered levels of relationship quality—remains unanswered and future studies would benefit from delving further into this gap in the literature. Due to the complex and ever-changing nature of relationships and life circumstances, it is most likely that the association between neuroticism and relationship quality could be bidirectional and cyclical in nature- that

is, that as neurotic emotions, behaviors, and interpretations increase, relationship quality decreases. Or it could be that as relationship quality decreases, neurotic emotions, behaviors, and interpretations increase. At this point, there is no telling which comes first.

Clinical Applications

Given the negative association between neuroticism and relationship quality, there are obvious clinical applications. Couples' therapists would do well to routinely assess personality traits using an evidence-based measure like the NEO (McCrae & Costa, 1987) or the EPQ (Eysenck & Eysenck, 1993). This would be a particularly pertinent intervention if neuroticism is primary in temporal precedence, leading to subsequent lower levels of relationship quality. Performing such an assessment prior to treatment would tell clinician's where each partner's level of neuroticism is and help them tailor their treatments to be more specific for that couple. For example, if one or both partners are high in neuroticism, a behavioral approach given what we know and assert about how neuroticism operates in relationships could be to increase the amount of positive emotions, behaviors and interpretations in a relationship and decrease the amount of negative emotions, behaviors and interpretations in a relationship. Specifically, increasing positive emotions, behaviors, and interpretations look like a clinician setting goals with a couple to give each other the benefit of the doubt at least once in the coming week, and to do something nice for one another as a surprise (Braithwaite & Fincham, 2007). Decreasing negative emotions, behaviors, and interpretations could look like taking a "time out" from a potential argument until one or both partners have "cooled off" and can approach the discussion rationally (Braithwaite & Fincham, 2007). Further, some studies have also found the sexual relationship to be a mediator in how neuroticism affects relationship quality (Fisher & McNulty,

2008; Russell & McNulty, 2011). This could be another level for intervention for a couples' therapist to improve the relationship functioning of individuals with neuroticism.

On the other hand, as we do not know which variable comes first (lower levels of relationship quality leading to higher levels of neuroticism or higher levels of neuroticism leading to lower levels of relationship quality), another appropriate level of intervention could be to simply increase the relationship quality. Couples therapists could help partners to meet one another's needs in a more targeted way, improve communication between the couple, or could help improve the sexual relationship, as these interventions typically see increases in relationship quality (Dewitte, & Mayer, 2018; Yoo et al., 2014). If it happens that higher levels of relationship quality decrease effects of neuroticism, then couples should still see gains in their relationship by simply targeting the relationship itself.

Strengths and Limitations

An obvious strength of this meta-analysis is the vast number of studies used to aggregate our overall effect size between neuroticism and relationship quality (151 studies overall, with 160 independent effect sizes). With a sample of studies this large, I have been able to truly test the limits of this relationship by seeing if the same association exists among different cultures, countries, ages, lengths of relationship, and many other factors. I have also been able to test the association across different ways of measuring, different ways of reporting (actor and partner), different countries, and across time. Even with such a large sample of studies, I was able to establish that the bivariate association between the two variables fell within a narrow range centered about -.22.

One limitation in the current study, but that is reflective of the body of literature at large, is the lack of studies that gather information on how neuroticism and relationship quality are

associated among LGBTQIA+ couples. Only a handful of studies have looked into these particular types of relationships, and the field would do well to continue gathering data among this population. Another limitation of this study is its inability to fully comment on a case for causation between heightened neuroticism and lower levels of relationship quality due to the subjectivity and ethical considerations of some of Hill's criteria for causation (Hill, 1965). Despite this, I have tried to objectively measure and outline a case for causation as best I can. There are obvious ways that the field will never be able to meet the entire criteria for causation, as institutional governing bodies would likely reject proposals to use an experimental design, in which some subjects would need to "act with more neurotic behaviors" so that experimenters can see what happens to a relationship. With that in mind, there are obvious limitations to what can be done, but I have tried our best to feasibly meet all the criteria that I can.

To conclude, this meta-analysis was able to thoroughly assess the relationship between neuroticism and relationship quality in light of Hill's criteria for causation. I considered every study I could find (written in English) that has already explored this relationship. By doing so, I had an extremely large sample of participants from diverse backgrounds and with diverse relationships, allowing us to test the limits of this highly replicated relationship and comment on a possible case for causation. I have confirmed with meta-analytic methods the common finding that lowered levels of relationship quality is associated with higher levels of neuroticism, and the bi-variate association exists within a narrow range centered at about -.22. Also, I have extended the current knowledge by providing: 1) a model for how neuroticism operates, 2) whether criteria for causation has been met, and 3) broad clinical applications for how a clinician may approach couples who are struggling in their relationship due to higher levels of neuroticism. Each of these extensions have provided ideas for future studies that will allow us to more fully understand how

neuroticism operates and what further needs to be done to allow all the criteria for causation to be met. These outcomes can provide the field with a basis for future studies that can inform our knowledge further on the complex relationship between neuroticism and relationship quality.

References

Adams, J. M., & Jones, W. H. (1997). The conceptualization of marital commitment: An integrative analysis. *Journal of Personality and Social Psychology*, *72*(5), 1177.

Adler, A. (1924) *The practice and theory of individual psychology.* New York, Harcourt, Brace, Jovanovich

Alexandrov, E. O., Cowan, P. A., & Cowan, C. (2005). Couple attachment and the quality of marital relationships: Method and concept in the validation of the new couple attachment interview and coding system. *Attachment & Human Development*, 7(2), 123–152. https://doi.org/10.1080/14616730500155170

Almiro, P. A., Moura, O., & Simões, M. R. (2016). Psychometric properties of the European Portuguese version of the Eysenck Personality Questionnaire—Revised (EPQ-R). *Personality and Individual Differences*, *88*, 88–93. https://doi.org/10.1016/j.paid.2015.08.050

Amato, P. R. (1999). Children of divorced parents as young adults. In E. M. Hetherington (Ed.), *Coping with divorce, single parenting, and remarriage: A risk and resiliency perspective* (p. 147–163). Lawrence Erlbaum Associates Publishers.

Aníbal González-Rivera, J. (2020). Reliability and validity of the Relationship Assessment Scale in Puerto Rican women. *Informes Psicológicos*, *20*(1), 11–18. https://doi.org/10.18566/infpsic.v20n1a01

Atkinson, M., & Violato, C. (1994). Neuroticism and coping with anger: The trans-situational consistency of coping responses. *Personality and Individual Differences*, *17*(6), 769–782.

Auersperg, F., Vlasak, T., Ponocny, I., & Barth, A. (2019). Long-term effects of parental divorce on mental health – A meta-analysis. *Journal of Psychiatric Research*, *119*, 107–115. https://doi.org/10.1016/j.jpsychires.2019.09.011

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*(2), 191.

Beach, S. R., Katz, J., Kim, S., & Brody, G. H. (2003). Prospective effects of marital satisfaction on depressive symptoms in established marriages: A dyadic model. *Journal of Social and Personal Relationships*, *20*(3), 355–371.

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbauch, J. (1961). Beck Depression Inventory. *PsycTESTS.*

Beck, A. T., Epstein, N., Brown, G., & Steer, R. (1988). Beck Anxiety Inventory. *PsycTESTS.*

Beck, A. T., Steer, R. A., Ball, R., & Ranieri, W. F. (1996). Comparison of Beck Depression Inventories-IA and-II in psychiatric outpatients. *Journal of Personality Assessment*, *67*(3), 588-597.

Bergner, R. M. (2020). What is personality? Two myths and a definition. *New Ideas in Psychology, 57.*

Bianchi, R. (2018). Burnout is more strongly linked to neuroticism than to work-contextualized factors. *Psychiatry Research*, *270*, 901–905. https://doi.org/10.1016/j.psychres.2018.11.015

Billstedt, E., Waern, M., Falk, H., Duberstein, P., Östling, S., Hällström, T., & Skoog, I. (2017). Time trends in Murray's psychogenic needs over three decades in Swedish 75-year-olds. *Gerontology*, *63*(1), 45–54.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis.* Wiley.

Borenstein, M. (2019). *Common mistakes in meta-analysis and how to avoid them*. Biostat, Incorporated.

Boulard, A., Gauthier, J.-M., & Born, M. (2014). Center for Epidemiologic Studies Depression Scale—French Version. *PsycTESTS*.

Boyle, G. J., Stankov, L., Martin, N. G., Petrides, K. V., Eysenck, M. W., & Ortet, G. (2016). Hans J. Eysenck and Raymond B. Cattell on intelligence and personality. *Personality & Individual Differences*, *103*, 40–47. https://doi.org/10.1016/j.paid.2016.04.029

Bradbury, T. N., & Fincham, F. D. (1991). A contextual model for advancing the study of marital interaction. *Cognition in Close Relationships*, 127–147.

Bradbury, T. N., Fincham, F. D., & Beach, S. H. (2000). Research on the nature and determinants of marital satisfaction: A decade in review. *Journal of Marriage & Family*, 62(4), 964–980.

Bradbury, T. N., & Karney, B. R. (2013). *Intimate relationships*. (2nd ed.). Norton.

Braithwaite, S. R., & Fincham, F. D. (2007). ePREP: Computer based prevention of relationship dysfunction, depression and anxiety. *Journal of Social and Clinical Psychology*, *26*(5), 609–622.

Braithwaite, S. R., Mitchell, C. M., Selby, E. A., & Fincham, F. D. (2016). Trait forgiveness and enduring vulnerabilities: Neuroticism and catastrophizing influence relationship satisfaction via less forgiveness. *Personality and Individual Differences, 94*, 237–246.

Branchaud, M. (2019). Attachment styles' role in marital satisfaction and divorce [ProQuest Information & Learning]. In *Dissertation Abstracts International: Section B: The Sciences and Engineering* (Vol. 80, Issue 4-B(E)).

Byrom, N. C., & Murphy, R. A. (2013). It's not just what we encode, but how we encode it: Associations between neuroticism and learning. *Journal of Personality*, *81*(4), 345–354. https://doi.org/10.1111/jopy.12022

Calmes, C. A., & Roberts, J. E. (2008). Rumination in interpersonal relationships: Does co-rumination explain gender differences in emotional distress and relationship satisfaction among college students? *Cognitive Therapy and Research*, *32*(4), 577–590.

Caprara, G. V. (1992). Reflections on the recent history and the present challenges of personality psychology. *European Journal of Personality*, *6*(5), 345–358. https://doi.org/10.1002/per.2410060503

Card, N. A. (2015). *Applied meta-analysis for social science research*. Guilford Publications.

Chen, Z., Tanaka, N., Uji, M., Hiramura, H., Shikai, N., Fujihara, S., & Kitamura, T. (2007). The role of personalities in the marital adjustment of Japanese couples. *Social Behavior and Personality: An International Journal*, *35*(4), 561–572.

Chun, S.-Y., Jang, S.-Y., Choi, J.-W., Shin, J., & Park, E.-C. (2016). Long-term effects of parental divorce timing on depression: A population-based longitudinal study. *International Journal of Social Psychiatry*, *62*(7), 645–650.

Costa, P. T., & McCrae, R. R. (1980). Influence of extraversion and neuroticism on subjective well-being: Happy and unhappy people. *Journal of Personality and Social Psychology*, *38*(4), 668.

Costa Jr, P. T., McCrae, R. R., Zonderman, A. B., Barbano, H. E., Lebowitz, B., & Larson, D. M. (1986). Cross-sectional studies of personality in a national sample: II. Stability in neuroticism, extraversion, and openness. *Psychology and Aging, 1*(2), 144.

Costa Jr, P. T., & McCrae, R. R. (1988). From catalog to classification: Murray's needs and the five-factor model. *Journal of Personality and Social Psychology*, *55*(2), 258–265. https://doi.org/10.1037/0022-3514.55.2.258

Costa Jr, P. T., & McCrae, R. R. (1992). Four ways five factors are basic. *Personality and Individual Differences*, *13*(6), 653–665.

Costa Jr, P. T., & McCrae, R. R. (1992). Revised NEO Personality Inventory. *PsycTESTS*.

Costa Jr, P. T., & McCrae, R. R. (2017). The NEO Inventories as instruments of psychological theory. *The Oxford handbook of the five factor model*, 11-37.

Cote´, S. & Moskowitz, D. S. (1998). On the dynamic covariation between interpersonal behavior and affect: Prediction from Neuroticism, extraversion, and agreeableness. *Journal of Personality and Social Psychology*, 75, 1032–1046.

Crandall, A., Powell, E. A., Bradford, G. C., Magnusson, B. M., Hanson, C. L., Barnes, M. D., Novilla, M. L. B., & Bean, R. A. (2019). Maslow's hierarchy of needs as a framework for understanding adolescent depressive symptoms over time. *Journal of Child and Family Studies*, *29*(2), 273–281. https://doi.org/10.1007/s10826-019-01577-4

Cronbach, L., J. (1970) *Essentials of psychological testing (3rd Ed)*. New York: Harper & Row.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*(4), 281.

Cundiff, J. M., Smith, T. W., & Frandsen, C. A. (2012). Incremental validity of spouse ratings versus self-reports of personality as predictors of marital quality and behavior during marital conflict. *Psychological Assessment*, *24*(3), 676.

Dewitte, M., & Mayer, A. (2018). Exploring the link between daily relationship quality, sexual desire, and sexual activity in couples. *Archives of Sexual Behavior*, *47*(6), 1675–1686.

Decuyper, M., De Bolle, M., & De Fruyt, F. (2012). Personality similarity, perceptual accuracy, and relationship satisfaction in dating and married couples. *Personal Relationships*, *19*(1), 128-145.

Diefenbach, H., & Opp, K. D. (2007). When and why do people think there should be a divorce? An application of the factorial survey. *Rationality and Society*, *19*(4), 485–517.

Dumont, F. (2010). *A history of personality psychology: Theory, science, and research from Hellenism to the twenty-first century*. Cambridge University Press.

Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot–based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, *56*(2), 455–463.

Elkatawneh, D. H. (2013). Freud's Psycho-Sexual Stages of Development. *Available at SSRN 2364215*.

Eysenck, H. J. (1970). *The structure of human personality (3rd Ed)*. Mouthen & Co., London.

Eysenck, H. J. (1992). Four ways five factors are not basic. *Personality and Individual Differences*, *13*(6), 667–673. https://doi.org/10.1016/0191-8869(92)90237-J

Eysenck, S., & Barrett, P. (2013). Re-introduction to cross-cultural studies of the EPQ. *Personality and Individual Differences*, *54*(4), 485–489.

Eysenck, H. J., & Eysenck, S. B. G. (1993). Eysenck Personality Questionnaire-Revised. *PsycTESTS.*

Fisher, T. D., & McNulty, J. K. (2008). Neuroticism and marital satisfaction: The mediating role played by the sexual relationship. *Journal of Family Psychology*, *22*(1), 112.

Forrester, J. (1991). *The seductions of psychoanalysis: Freud, Lacan and Derrida* (Vol. 26). Cambridge University Press.

Freeston, M. H., & Pléchaty, M. (1997). Reconsiderations of the Locke-Wallace Marital Adjustment Test: Is it still relevant for the 1990s? *Psychological Reports*, *81*(2), 419–434. https://doi.org/10.2466/PR0.81.6.419–434

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

Geen, R. G., Beatty, W. W., & Arkin, R. M. (1984). *Human motivation: Physiological, behavioral, and social approaches*. Allyn & Bacon.

Gentry, T. A., Wakefield Jr., J. A., & Friedman, A. F. (1985). MMPI Scales for measuring Eysenck's Personality Factors. *Journal of Personality Assessment*, *49*(2), 146. https://doi.org/10.1207/s15327752jpa4902_7

Givertz, M., Segrin, C., & Hanzal, A. (2009). The association between satisfaction and commitment differs across marital couple types. *Communication Research*, *36*(4), 561–584. https://doi.org/10.1177/2F0093650209333035

Goldberg, L. R. (1990). An alternative" description of personality": The big-five factor structure. *Journal of Personality and Social Psychology*, *59*(6), 1216.

Goldberg, L. R., & Rosolack, T. K. (1994). The Big Five factor structure as an integrative framework: An empirical comparison with Eysenck's PEN model. *The developing structure of temperament and personality from infancy to adulthood*, 7–35.

Gottman, J. M. (1990). How marriages change. *Depression and Aggression in Family Interaction*, *75*, 101.

Grames, H. A., Miller, R. B., Robinson, W. D., Higgins, D. J., & Hinton, W. J. (2008). A test of contextual theory: The relationship among relational ethics, marital satisfaction, health problems, and depression. *Contemporary Family Therapy*, *30*(4), 183-198.

Gul, L., & Jahangir, S. F. (2019). The effectiveness of mindfulness-based stress reduction programme (MBSRP) and Sufi Meditation (SM) in the treatment of neurotic anxiety among females. *FWU Journal of Social Sciences*, *13*(1).

Hall, C. S., & Lindzey, G. (1985). *Introduction to theories of personality.* John Wiley & Sons

Hendrick, S. S. (1988). A generic measure of relationship satisfaction. *Journal of Marriage and the Family*, *50*(1), 93–98.

Hill A. B. (1965). The environment and disease: Association or causation? *Proceedings of the Royal Society of Medicine*, *58,* 295.

Horney, K. (2013). *New ways in psychoanalysis*. Routledge.

Huston, T. L., & Burgess, R. L. (1979). Social exchange in developing relationships: An overview. In R. L. Burgess & T. L. Huston (Eds.), *Social exchange in developing relationships* (pp. 3–28). New York: Academic Press.

Iveniuk, J., Waite, L. J., Laumann, E., McClintock, M. K., & Tiedt, A. D. (2014). Marital conflict in older couples: Positivity, personality, and health. *Journal of Marriage and the Family*, 76, 130–144. https://doi.org/10.1111/jomf.12085

Jang, K. L., Livesley, W. J., & Vemon, P. A. (1996). Heritability of the Big Five Personality dimensions and their facets: A twin study. *Journal of Personality*, *64*(3), 577-592.

John, O. P., Donahue, E. M., & Kentle, R. L. (1991). Big Five Inventory. *PsycTESTS*.

John, O. P., & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. *Handbook of Personality: Theory and Research*, *2*(1999), 102–138.

Johnson, M. P. (1991), Commitment to personal relationships. *Advances in Personal Relationships*, 117–143.

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, 1091–1108.

Karney, B. R., & Bradbury, T. N. (1997). Neuroticism, marital interaction, and the trajectory of marital satisfaction. *Journal of Personality and Social Psychology*, *72*(5), 1075.

Karney, B. R., Bradbury, T. N., Fincham, F. D., & Sullivan, K. T. (1994). The role of negative affectivity in the association between attributions and marital satisfaction. *Journal of Personality and Social Psychology*, *66*(2), 413.

Kean, J., & Reilly, J. (2014). Item response theory. *Handbook for clinical research: Design, statistics and implementation*, 195–198.

Kelley, H. H., & Thibaut, J. W. (1978). *Interpersonal relations: A theory of interdependence*. New York: Wiley.

Kendler, K. S., Gatz, M., Gardner, C. O., & Pedersen, N. L. (2006). Personality and major depression: A Swedish longitudinal, population-based twin study. *Archives of General Psychiatry*, *63*(10), 1113–1120.

Lahey, B. B. (2009). Public health significance of neuroticism. *American Psychologist, 64*(4), 241–256. https://doi.org/10.1037/a0015309

Laub, J. H., Nagin, D. S., & Sampson, R. J. (2017). Trajectories of change in criminal offending: Good marriages and the desistance process. In *The termination of criminal careers* (pp. 433-446). Routledge.

Lavner, J. A., & Bradbury, T. N. (2012). Why do even satisfied newlyweds eventually go on to divorce? *Journal of Family Psychology*, *26*(1), 1.

Lester, D., Hvezda, J., Sullivan, S., & Plourde, R. (1983). Maslow's hierarchy of needs and psychological health. *The Journal of General Psychology*, *109*(1), 83–85.

Locke, H. J., & Wallace, K. M. (1959). Marital Adjustment Test. *PsycTESTS.*

Lundy, A. (1988) Instructional set and Thematic Apperception Test validity. *Journal of Personality Assessment, 52*(2), 309.

Magalhães, E.. Salgueira, A., Gonzalez, A.-J., Costa, J. J., Costa, M. J., Costa, P., & de Lima, M. P. (2014) Neo Five-Factor Inventory – Portuguese version. *PsycTESTS.*

Malouff, J.M., Thorsteinsson, E.B., & Schutte, N.S. (2005). The relationship between the five-factor model of personality and symptoms of clinical disorders: A meta-analysis. *Journal of Psychopathology and Behavioral Assessment*, 27, 101–114.

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

Marsh, H. W., Nagengast, B., & Morin, A. J. (2013). Measurement invariance of big-five factors over the life span: ESEM tests of gender, age, plasticity, maturity, and la dolce vita effects. *Developmental Psychology*, *49*(6), 1194.

McAdams, D. P. (1997). A conceptual history of personality psychology. *Handbook of personality psychology.* (pp. 3–39). Academic Press.

McAdams, D. P., & Pals, J. L. (2006). A new Big Five: Fundamental principles for an integrative science of personality. *American Psychologist*, *61*(3), 204–217. https://doi.org/10.1037/0003-066X.61.3.204

McCrae, R. R., & Costa, P. T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology, 52*(1), 81–90. https://doi.org/10.1037/0022-3514.52.1.81

McCrae, R. R., & Costa Jr, P. T. (2004a). A contemplated revision of the NEO Five-Factor Inventory. *Personality and Individual Differences, 36*(3), 587–596.

McCrae, R. R., & Costa, P. T., Jr. (2004b). NEO Five-Factor Inventory--Revised. *PsycTESTS*. https://doi.org/10.1037/t07553-000

McNulty, J. K. (2008). Neuroticism and interpersonal negativity: The independent contributions of perceptions and behaviors. *Personality and Social Psychology Bulletin*, *34*(11), 1439–1450.

Meyers, S. A., & Landsberger, S. A. (2002). Direct and indirect pathways between adult attachment style and marital satisfaction. *Personal Relationships*, *9*(2), 159–172. https://doi.org/10.1111/1475-6811.00010

Miletic, M. P. (2002). The introduction of a feminine psychology to psychoanalysis: Karen Horney's legacy. *Contemporary Psychoanalysis*, *38*(2), 287–299.

Miller, R. B., Anderson, S., & Keala, D. K. (2004). Is Bowen theory valid? A review of basic research. *Journal of Marital and Family Therapy*, *30*(4), 453–66.

Miller, R. B., Mason, T. M., Canlas, J. M., Wang, D., Nelson, D. A., & Hart, C. H. (2013). Marital satisfaction and depressive symptoms in China. *Journal of Family Psychology*, *27*(4), 677.

Mischel, W., Shoda, Y., & Ayduk, O. (2007). *Introduction to personality: Toward an integrative science of the person*. John Wiley & Sons.

Moini, S. (2017). What predicts marital satisfaction? Defining marital satisfaction through attachment and differentiation [ProQuest Information & Learning]. In *Dissertation Abstracts International: Section B: The Sciences and Engineering 77*(8–B(E)).

Mund, M., Finn, C., Hagemeyer, B., & Neyer, F. J. (2016). Understanding dynamic transactions between personality traits and partner relationships. *Current Directions in Psychological Science, 25*, 411–416. https://doi.org/10.1177/2F0963721416659458

Murstein, B. (1963) *Theory and research in projective techniques (emphasizing the TAT).* New York: Wiley

Nakonezny, P. A., & Denton, W. H. (2008). Marital relationships: A social exchange theory perspective. *The American Journal of Family Therapy*, *36*(5), 402–412.

Naragon-Gainey, K., & Watson, D. (2018). What lies beyond neuroticism? An examination of the unique contributions of social-cognitive vulnerabilities to internalizing disorders. *Assessment*, *25*(2), 143–158. https://doi.org/10.1177/1073191116659741

Netter, P., Hennig, J., & Munk, A. J. (2020). Principles and approaches in Hans Eysenck's personality theory: Their renaissance and development in current neurochemical research on individual differences. *Personality and Individual Differences*, 109975.

Nordhagen, T. (2001). *Beck Anxiety Inventory: translation and validation of a Norwegian version* (Master's thesis, The University of Bergen).

Norton, R. (1983). Quality Marriage Index. *PsycTESTS.*

O'Rourke, N. (2005). Personality, cognitive adaptation, and marital satisfaction as predictors of well-being among older married adults. *Canadian Journal on Aging/La Revue Canadienne du Vieillissement*, *24*(3), 211–224.

Perkmen, S., Toy, S., Caracuel, A., & Shelley, M. (2018). Cross-cultural search for Big Five: Development of a scale to compare personality traits of pre-service elementary school teachers in Turkey and Spain. *Asia Pacific Education Review*, *19*(4), 459–468. https://doi.org/10.1007/s12564-018-9549-2

Peterson, R. A., & Brown, S. P. (2005). On the use of beta coefficients in meta analysis. *Journal of Applied Psychology*, *90*(1), 175.

Pryor, J., & Roberts, J. (2005). What is commitment? How married and cohabiting parents talk about their relationships. *Family Matters*, *71*, 24–31.

Qing Zeng, P., Draper, T. W., & Chongde Lin. (2003). Personality structure in Chinese school children: Is the Five-Factor Model adequate? *African & Asian Studies*, *2*(3), 213–232.

Radloff, L. S. (1977). Center for Epidemiological Studies Depression Scale. *PsycTESTS.*

Rafienia, P., Azadfallah, P., Fathi-Ashtiani, A., & Rasoulzadeh-Tabatabaiei, K. (2008). The role of extraversion, neuroticism and positive and negative mood in emotional information processing. *Personality and Individual Differences*, *44*(2), 392–402.

Rammstedt, B., Danner, D., Soto, C. J., & John, O. P. (2020). Validation of the short and extra-short forms of the Big Five Inventory-2 (BFI-2) and their German adaptations. *European Journal of Psychological Assessment*, *36*(1), 149–161. https://doi.org/10.1027/1015-5759/a000481 (Supplemental)

Razeghi, N., Nikiju, M., Mujembari, A. K., & Masihi, A. Z. (2011). Relationship between Big Five personality factors and marital satisfaction. *Journal of Iranian Psychologists*, *7*(27), 269–278.

Ricoeur, P. (1992). The question of proof in Freud's psychoanalytic writings. In R. B. Miller (Ed.), *The restoration of dialogue: Readings in the philosophy of clinical psychology* (pp. 347–365). American Psychological Association.

Ripley, J. S., Worthington, E. L., Garthe, R. C., Davis, D. E., Hook, J. N., Reid, C. A., Van Tongeren, D. R., Voltmer, A., Nonterah, C. W., Cowden, R. G., Coetzer-Liversage, A., Cairo, A., Joynt, S., & Akpalu, B. (2018). Trait forgiveness and dyadic adjustment predict postnatal depression. *Journal of Child & Family Studies*, *27*(7), 2185–2192. https://doi.org/10.1007/s10826-018-1053-0

Roberson, P. N. E., & Fincham, F. (2018). Is relationship quality linked to diabetes risk and management?: It depends on what you look at. *Families, Systems, & Health*, *36*(3), 315–326. https://doi.org/10.1037/fsh0000336

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.

Rogers, C. R. (2013). A theory of therapy and personality change: As developed in the client-centered framework". *Perspectives in Abnormal Behavior: Pergamon General Psychology Series*, *341*.

Russell, V. M., & McNulty, J. K. (2011). Frequent sex protects intimates from the negative implications of their neuroticism. *Social Psychological and Personality Science*, *2*(2), 220–227.

Russell, R. J. H., & Wells, P. A. (1994a). Personality and quality of marriage. *British Journal of Psychology*, *85*, 161–168.

Russell, R. J., & Wells, P. A. (1994b). Predictors of happiness in married couples. *Personality and Individual Differences*, *17*(3), 313–321

Sands, A., Thompson, E. J., & Gaysina, D. (2017). Long-term influences of parental divorce on offspring affective disorders: A systematic review and meta-analysis. *Journal of Affective Disorders*, *218*, 105–114. https://doi.org/10.1016/j.jad.2017.04.015

Sayehmiri, K., Kareem, K. I., Abdi, K., Dalvand, S., & Gheshlagh, R. G. (2020). The relationship between personality traits and marital satisfaction: A systematic review and meta-analysis. *BMC Psychology*, *8*(1), 15.

Schaffhuser, K., Wagner, J., Lüdtke, O., & Allemand, M. (2014). Dyadic longitudinal interplay between personality and relationship satisfaction: A focus on neuroticism and self-esteem. *Journal of Research in Personality*, *53*, 124–133.

Schultz, D.P., & Schultz, S.E. (2017). *Theories of personality*. Australia: Cengage Learning.

Scollon, C. N., & Diener, E. (2006). Love, work, and changes in extraversion and neuroticism over time. *Journal of Personality and Social Psychology*, *91*(6), 1152.

Sharma, M. K., & Raju, M. (2013). Relationship of personality dimensions and aggression in romantic relationship among youth. *Indian Journal of Psychological Medicine*, *35*(2), 197-202.

Shen, J., Brdiczka, O., & Liu, J. (2015). A study of Facebook behavior: What does it tell about your neuroticism and extraversion? *Computers in Human Behavior*, *45*, 32–38. https://doi.org/10.1016/j.chb.2014.11.067

Sherman, M. (1941). Theories of personality. In *Basic problems of behavior.* (pp. 108–158). Longmans, Green and Co.

Shoda, Y., LeeTiernan, S., & Mischel, W. (2002). Personality as a dynamical system: Emergence of stability and distinctiveness from intra and interpersonal interactions. *Personality and Social Psychology Review*, *6*(4), 316-325.

Skinner, B. F. (1981). Selection by consequences. *Science*, 213, 501–504. https://doi.org/10.1017/S0140525X0002673X

Skowron, E. A. (2000). The role of differentiation of self in marital adjustment. *Journal of Counseling Psychology*, 47(2), 229-237. https://doi.org/10.1037/0022-0167.47.2.229

Solomon, B. C., & Jackson, J. J. (2014). Why do personality traits predict divorce? Multiple pathways through satisfaction. *Journal of Personality and Social Psychology*, *106*(6), 978–996. https://doi.org/10.1037/a0036190 (Supplemental)

Soto, C. J., & John, O. P. (2017). Big Five Inventory-2. *PsycTESTS.*

Soto, C. J., Kronauer, A., & Liang, J. K. (2015). Five-factor model of personality. *The encyclopedia of adulthood and aging*, 1–5.

Spanier, G. B. (1976) Dyadic Adjustment Scale. *PsycTESTS.*

Stack, S., & Eshleman, J. R. (1998). Marital status and happiness: A 17-nation study. *Journal of Marriage and the Family*, 527-536.

Stanley, S. M., Carlson, R., Rhoades, G. K., Markman, H. J., Ritchie, L., & Hawkins, A. J. (2020). Best practices in relationship education. *Family Relations, 69,* 497–519. https://doi.org/10.1111/fare.12419

Stauffer, M. (2015). Inducing self-compassion and the impact on neuroticism and other personality characteristics. *USC Aiken Psychology Theses*. 29.

Teasdale, J. D., & Rachman, S. (1983). Cognitions and mood: Clinical aspects and applications. *Advances in Behaviour Research and Therapy*, *5*(1), 88.

Thibaut, J. W., & Kelley, H. H.(1959). *The social psychology of groups.* New York: Wiley.

Tong, W., Li, P., Zhou, N., He, Q., Ju, X., Lan, J., ... & Fang, X. (2018). Marriage improves neuroticism in Chinese newlyweds: Communication and marital affect as mediators. *Journal of Family Psychology*, *32*(7), 986.

Törnroos, M., Jokela, M., & Hakulinen, C. (2019). The relationship between personality and job satisfaction across occupations. *Personality and Individual Differences*, *145*, 82–88. https://doi.org/10.1016/j.paid.2019.03.027

Vázquez, F. L., Otero, P., López, L., Blanco, V., José Ferraces, M., & Torres, Á. (2019). Eysenck Personality Questionnaire Revised-Abbreviated for informal caregivers. *Revista Iberoamericana De Psicología Y Salud*, *10*(2), 90–106.

Vezzetti, V. C. (2016). New approaches to divorce with children: A problem of public health. *Health Psychology Open*, *3*(2), 2055102916678105. https://doi.org/10.1177/2F2055102916678105

Waite, L. J., & Gallagher, M. (2001). The case for marriage: Why married people are happier, healthier, and better off financially. *Random House Digital, Inc.*

Watson, D., & Clark, L. A. (1984). Negative affectivity: The disposition to experience aversive emotional states. *Psychological Bulletin*, 96, 465–490.

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.

Whitton, S. W., & Kuryluk, A. D. (2012). Relationship satisfaction and depressive symptoms in emerging adults: Cross-sectional associations and moderating effects of relationship characteristics. *Journal of Family Psychology*, *26*(2), 226.

Whitton, S. W., & Kuryluk, A. D. (2014). Associations between relationship quality and depressive symptoms in same-sex couples. *Journal of Family Psychology*, *28*(4), 571–576. https://doi.org/10.1037/fam0000011

Whitton, S. W., Stanley, S. M., Markman, H. J., & Johnson, C. A. (2013). Attitudes toward divorce, commitment, and divorce proneness in first marriages and remarriages. *Journal of Marriage and Family*, *75*(2), 276–287. https://doi.org/10.1111/jomf.12008

Wolfinger, N. H. (2000). Beyond the intergenerational transmission of divorce: Do people replicate the patterns of marital instability they grew up with? *Journal of Family Issues*, *21*(8), 1061–1086.

Yoo, H., Bartle-Haring, S., Day, R. D., & Gangamma, R. (2014). Couple communication, emotional and sexual intimacy, and relationship satisfaction. *Journal of Sex & Marital Therapy*, *40*(4), 275–293.

Yoon, K. L., Maltby, J., & Joormann, J. (2013). A pathway from neuroticism to depression: Examining the role of emotion regulation. *Anxiety, Stress & Coping*, *26*(5), 558–572. https://doi.org/10.1080/10615806.2012.734810

Zheng, L., Goldberg, L. R., Zheng, Y, Zhao, Y., Tang, Y., & Liu, L. (2008). Reliability and concurrent validity of the IPIP Big-Five factor markers in China: Consistencies in factor structure between internet-obtained heterosexual and homosexual samples. *Personality & Individual Differences, 45*(7), 649–654.

Zhong, M., He, Q., Lan, J., Ju, X., Li, X., & Fang, X. (2016). Marital quality and marital stability in newlywed couples: Mediating of marital commitment. *Chinese Journal of Clinical Psychology*, *24*(6), 1064–1068.

Zimmermann, T., de Zwaan, M., & Heinrichs, N. (2019). The German version of the Quality of Marriage Index: Psychometric properties in a representative sample and population-based norms. *PLoS ONE*, *14*(2), 1–11. https://doi.org/10.1371/journal.pone.0212758

Zvolensky, M. J., Shepherd, J. M., Garey, L., Case, K., & Gallagher, M. W. (2020). The influence of neuroticism in terms of E-cigarette dependence and beliefs about use and quitting among dual users of combustible and electronic cigarettes. *Addictive Behaviors*, *107*. https://doi.org/10.1016/j.addbeh.2020.106396

Appendix A

Data Entry Form Questions

1. What is your name?
2. First Author
3. Second Author
4. Third Author
5. Fourth Author
6. Fifth Author
7. Sixth Author and beyond
8. Year
9. Title
10. Source - Typically the title of the journal where the article was published. If a dissertation, write “Dissertation”.
11. Has the study been published in a peer-reviewed journal?
12. Do the authors provide information about funding? - If yes, select “Other” and write the name of the funding source.
13. Abstract - Copy and paste the abstract below.
14. Sample size for husbands - If they only report on couples, divide that in half to find sample size for husbands.
15. Sample size for wives - If they only report on couples, divide that in half to find sample size for wives.
16. From which country was the sample obtained? - If this is not explicitly stated in the manuscript, provide as much info as you can.
17. Mean Husband Age - Mean age of husbands in the sample. If separate means are not provided for husbands and wives, report the sample mean as the mean for both husbands and wives. If this information is not provided input "Not reported".
18. Standard Deviation of Husband Age - If separate values are not provided for husbands and wives, report the sample mean as the mean for both husbands and wives. If this information is not provided input "Not reported".

19. Mean Wife Age - Mean age of wives in the sample. If separate means are not provided for husbands and wives, report the sample mean as the mean for both husbands and wives. If this information is not provided input "Not reported".

20. Standard deviation of Wife Age - If separate values are not provided for husbands and wives, report the sample mean as the mean for both husbands and wives. If this information is not provided input "Not reported".

21. Percentage Sexual Minority - What percentage of respondents were not heterosexual? If this is not explicitly stated provide as much information as you can to help make an inference. For example, if they included only married couples, the study was published in 2005, and they recruited from Oklahoma we can assume they include 0% sexual minorities.

22. Percentage Racial Minority - If sample is from US or another predominantly White country what percentage of respondents were not White? If this information is not provided input "Not reported".

23. Average combined family income of sample - If this information is not provided input "Not reported".

24. Years of Education Guide

Primary school						High school						College/University				Graduate school						Post-graduate
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23+

25. Average years of education for husbands - If this information is not provided input "Not reported".

26. Average years of education for wives - If this information is not provided input "Not reported".

27. What percentage of couples had children? - If this information is not provided input "Not reported".

28. What was the average length of marriage for the sample? - If this information is not provided input "Not reported".

29. Unique characteristics of the sample - MTurk sample, college students, recruited from women's shelter, etc.

30. Was the data…

    a. Self-report of only one partner (e.g., one person from the couple reported on the relationship)

    b. Self-report from both partners (both partners reported on their own perceptions of satisfaction/neuroticism, etc.--dyadic data)

    c. Respondents reported on both their own and their spouses personality, behavior, etc. (as is often done with the CTS; dyadic data with self- and partner-report)

    d. Other:

31. What scale was used to measure neuroticism? - If more than one, report all and provide as much info as possible about when different scales were used

32. What was Cronbach's alpha for this scale? - Again, if more than one measure of the construct was used, provide info for all scales

33. Did this study provide information about relationship satisfaction?

    a. No

    b. Yes

    c. Other:

34. What scale was used to measure relationship satisfaction? - If more than one, report all and provide as much info as possible about when different scales were used. If this information is not provided input "Not reported".

35. What was Cronbach's Alpha for this scale? - Again, if more than one measure of the construct was used, provide info for all scales. If this information is not provided input "Not reported".

36. Did this study provide information about divorce? - County records, self-report, etc. Provide as much information as possible.

    a. No

    b. Yes

    c. Other:

37. How did they measure divorce? - County records, self-report, etc. Provide as much information as possible. If this information is not provided input "Not reported".

38. Is the study

    a. Cross-sectional

    b. Longitudinal

39. How many months between the assessment of neuroticism and relationship satisfaction - If the study is cross sectional, input 0

40. If the study is longitudinal, how many months between the assessment of neuroticism and divorce - If the study is cross sectional, input 0

41. If the study was longitudinal, what was the retention rate? - If the study was not longitudinal input "Not longitudinal".

42. Provide any information you can about missing data. - What approach they used (listwise deletion, FIML), evidence re: MCAR, MAR, MNAR

43. Are there any other characteristics about this study that would affect interpretations (e.g., it was interventional)

Actor Effects on Marital Satisfaction

44. Husband actor correlation between neuroticism and marital satisfaction - That is, the influence of husband neuroticism on his own marital satisfaction. If this information is not provided input "Not reported".

45. Sample size contributing to the effect size above - Do they report how many people are included in this correlation? If this information is provided, we need to know if it is different than the sample size listed previously. Look in tables and the results section for this (F or t values). If no more information is provided, input "No more information is provided".

46. Wife actor correlation between neuroticism and marital satisfaction - That is, the influence of wife neuroticism on her own marital satisfaction. If this information is not provided input "Not reported".

47. Sample size contributing to the effect size above - Do they report how many people are included in this correlation? If this information is provided, we need to know if it is different than the sample size listed previously. Look in tables and the results section for this (F or t values). If no more information is provided, input "No more information is provided".

Partner Effects on Marital Satisfaction

48. Husband partner correlation between neuroticism and marital satisfaction - That is, the influence of Husband neuroticism on wife marital satisfaction. If this information is not provided input "Not reported".

49. Sample size contributing to the effect size above - Do they report how many people are included in this correlation? If this information is provided, we need to know if it is different than the sample size listed previously. Look in tables and the results section for this (F or t values). If no more information is provided, input "No more information is provided".

50. Wife partner correlation between neuroticism and marital satisfaction - That is, the influence of Wife neuroticism on husband marital satisfaction. If this information is not provided input "Not reported".

51. Sample size contributing to the effect size above - Do they report how many people are included in this correlation? If this information is provided, we need to know if it is different than the sample size listed previously. Look in tables and the results section for this (F or t values). If no more information is provided, input "No more information is provided".

Correlation Between Husband Neuroticism and Divorce

52. Correlation between husband neuroticism and divorce - If this information is not provided input "Not reported".

53. What type of correlation did the authors report for the value above? - Pearsons r ? Point-biserial? If this information is not provided input "Not reported".

54. Sample size contributing to the effect size above - Do they report how many people are included in this correlation? If this information is provided, we need to know if it is different than the sample size listed previously. Look in tables and the results section for this (F or t values). If no more information is provided, input "No more information is provided".

55. Correlation between wife neuroticism and divorce - If this information is not provided input "Not reported".

56. What type of correlation did the authors report for the value above? - Pearsons r ? Point-biserial? If this information is not provided input "Not reported".

57. Sample size contributing to the effect size above - Do they report how many people are included in this correlation? If this information is provided, we need to know if it is different than the sample size listed previously. Look in tables and the results section for this (F or t values). If no more information is provided, input "No more information is provided".

Appendix B

Authors, sample size, country, measure, effect size, average relationship length (in years), average age of participants (in years), type of study

Study	N	Country	Measures	Effect Size	Rel Length	Age	Study Type
Altgelt, Reyes, French, Meltzer, & McNulty, 2018	216	USA	IPIP & QMI	-.161	0	29	LG
Amiri, Farhoodi, Abdolvand, Bidakhavidi, 2011	100	Iran	NEO-FFI & ENRICH	-.410	NR	NR	CS
Baker & McNulty, 2009	140	USA	EPQ & SD	-.217	1	26	CS
Barelds, 2005	1380	Netherlands	DPQ & DRQ	-.470	20	45	CS
Beach, Katz, Kim, & Brody, 2003	332	USA	CES-D & MAT	-.248	NR	40	LG
Beach & O'Leary, 1993	482	USA	BDI & MAT	-.399	1	24	LG
Bhagat & Hasan, 2014	200	India	EPQ & MSS	-.345	NR	39	CS
Bouchard & Arseneault, 2005	452	Canada	NEO-FFI & DAS	-.390	12	36	CS
Bouchard, Lussier, & Sabourn, 1999	446	Canada	NEO-FFI & DAS	-.242	9	35	CS
Braithwaite, Mitchell, Selby, & Fincham, 2015	355	USA	NEO-FFI & CSI-4	-.378	2	19	LG
Braithwaite, Mitchell, Selby, & Fincham, 2015	354	USA	NEO-FFI & CSI-4	-.300	2	19	LG
Brandt-Salmeri & Przybyła-Basista, 2019	97	Poland	BDI & GMQ	-.070	29	56	CS
Brock, Franz, & Ramsdell, 2020	318	USA	GTS & RQI	-.287	5	30	CS
Brudek, Steuden, & Jasik, 2018	120	Poland	NEO-PI-R & GMQ	-.236	NR	66	CS

Bruinsma, Peetoom, Millenaar, Köhler, Bakker, Koopmans, Pijnenburg, Verhey, & de Vugt, 2020	178	Netherlands	NEO-FFI & *Other	-.366	0	NR	LG
Buchanan, 2019	146	USA	NEO-FFI & CSI-16	-.350	NR	20	CS
Buckingham, Yamkovenko, Boring, Andrade, & lafolla, 2019	244	USA	**Other & QMI	-.240	4	36	CS
Bühler, Finkenauer, & Grob, 2020	326	Switzerland	BFI & RAS	-.210	24	50	LG
Calmes & Roberts, 2008	345	USA	BAI, BDI, CQ, & QRI	-.040	NR	20	CS
Cao, Yuan, Fine, Zhou, & Fang, 2019	536	China	NEO-FFI & QMI	-.209	1	29	LG
Cao, Zhou, Fang, & Fine, 2017	536	China	NEO-FFI & QMI	-.209	1	28	LG
Caughlin, Huston, & Houts, 2000	336	USA	16PF & MOQ	-.284	NR	23	LG
Chesnut, 2021	250	USA	3M-40, CSI-32, & CSI-4	-.045	27	41	CS
Chen, Tanaka, Uji, Hiramura, Shikai, Fujihara, & Kitamura, 2007	132	Japan	EPQ & SMAT	-.256			
Chopik & Lucas, 2019	5156	USA	BFI & ***Other	-.173	22	51	CS
Čikeš, Marić, & Šincek, 2018	196	Croatia	BFI & QMI	-.170	16	42	CS
Cirhinlioglu, Tepe, & Cirhinlioglu, 2016	976	Turkey	BFI & DAS	-.195	12	36	CS
Clark, Donnellan, & Robbins, 2020	900	USA	IPQ & ****Other	-.185	NR	38	LG
Claxton, O'Rourke, Smith, & DeLongis, 2012	50	USA/Canada	NEO-FFI & DAS	-.129	34	59	LG
Collins, 2009	344	USA	EPQ-N, MAT, & SD	-.256	4	27	LG
Cook, 2000	150	USA	NEO-PI & DAS	-.354	4	NR	LG
Cotter & Kerschner, 2018	419	USA	FCFFMQ & *****Other	-.130	9	37	CS
Cundiff, Smith, & Frandsen, 2012	600	USA	NEO-PI-R, CES-D, & MAT	-.308	27	54	CS

Daspe, Sabourin, Péloquin, Lussier, & Wright, 2013	944	Canada	NEO-FFI & DAS	-.164	13	41	CS
Davila, Karney, Hall, & Bradbury, 2003	344	USA	EPQ-N, BDI, MAT, & SD	-.291	0	24	LG
Decuyper, De Bolle, & De Fruyt, 2012	382	Belgium	NEO-PI-R & RAS	-.162	10	33	CS
Donnellan, Assad, Robins, & Conger, 2007	674	USA	MPQ-BF & QMI	-.206	NR	27	CS
Donnellan, Assad, Robins, & Conger, 2007	674	USA	MPQ-BF & QMI	-.280	NR	27	CS
Donnellan, Conger, & Bryant, 2004	836	USA	NEO-FFI & †Other	-.154	19	39	CS
Donnellan, Larsen-Rife, & Conger, 2005	580	USA	MPQ & QMI	-.320	2	18	LG
Drahman, Nubailah, & Yusof, 2018	110	Malaysia	BFI, DAS, & KMSS	.195	NR	NR	CS
Dyrenforth, 2011	11832	UK & Australia	BFI & †*Other	.136	21	52	CS
Eysenck & Wakefield, 1981	1132	USA	EPQ & MAT	-.234	9	35	CS
Eze, 2019	370	Nigeria	BFI & CSI	.804	NR	31	CS
Fani & Kheirabadi, 2011	161	Iran	NEO-FFI & DAS	-.377	NR	37	CS
Fincham, Beach, Harold, & Osborne, 1997	300	USA	BDI & MAT	-.368	0	27	LG
Finn, Mitte, & Neyer, 2013	210	Germany	NEO-FFI & RAS	-.130	7	30	CS
Fisher & McNulty, 2008	144	USA	BFI & SD	-.302	0	24	LG
Galinha, Garci-Martin, Oishi, Wirtz, & Esteves, 2016	497	India, Sweden, USA	PANAS & PWI	-.580	NR	19	CS

Galinha, Oishi, Pereira, Wirtz, & Esteves, 2013	1576	USA, Mozambique, Portugal	PANAS, BFI, SWLS, & PWI	-.640	NR	19	CS
Gana, Saada, Broc, Koleck, & Untas, 2017	396	France	HADS & DAS	-.292	14	37	CS
Gaunt, 2006	496	Israel	BABS & ENRICH	-.250	NR	31	CS
Grames, Miller, Robinson, Higgins, & Hinton, 2008	632	USA	CES-D & RDAS	-.391	14	44	CS
Gray, 2003	152	USA	BFI, PANAS-X, ORS, & QMI	-.070	7	35	CS
Gray, 2003	178	USA	BFI, PANAS-X, ORS, & QMI	-.375	1	22	CS
Großmann, 2017	192	Germany	YOUME-List & Das	-.530	3	NR	LG
Großmann, Hottung, & Krohn-Grimberghe, 2019	192	Germany	PD-I, MSI-R, DAS, & QPD	-.320	3	NR	LG
Herrick, 2018	294	USA	BDI & DAS	-.170	NR	33	CS
Hoppmann, & Blanchard-Fields	98	USA	NEO-FFI & RAS	-.330	47	72	CS
Jiang, Tang, He, Lin, & Zhou, 2021	1134	China	NEO-FFI & DAS	-.277	NR	36	CS
Jones, 2003	134	USA	NEO-PI, MAT, MSI-R, & DAS	-.270	10	41	CS
Karney & Bradbury, 1997	108	USA	EPQ-N, MAT, QMI, KMS, & SD	-.257	NR	24	CS
Karney, Bradbury, Fincham, & Sullivan, 1994	160	USA	BDI, EPQ-N, MAT, QMI, KMS, & SD	-.259	7	32	CS
Kelly & Conley, 1987	600	USA	PRS & †**Other	-.279	NR	NR	LG
Klymova & Dornisch, 2018	100	USA	IPIP-NEO-120 & RAS	-.112	NR	24	CS
Kurdek, 1997	516	USA	NEO-FFI, MDRCI, SLS, & CRSI	-.470	9	40	CS

Kurdek, 1998	396	USA	SCL-90-R, & DAS	-.243	NR	24	LG
Lavee & Ben-Ari, 2004	394	Israel	EPQ-R & EMS	-.221	17	42	CS
Lavner & Bradbury, 2012	272	USA	EPQ-N & MAT	-.235	0	27	LG
Lavner, Weiss, Miller, & Karney, 2018	338	USA	IPIP & QMI	-.191	0	24	LG
Lee, Rogge, & Reis, 2010	100	USA	EPQ-N, CSI, & MAT	-.260	2	23	LG
Lee, Taylor, Holbert, & Graham, 2019	1478	USA	MMPI-2-RF & DAS	-.030	18	41	CS
Leggett, Beachkofsky, & Leggett, 2014	102	USA	IPIP & DAS-4	-.520	15	42	CS
Leikas, Ilmarinen, Verkasalo, Vartiainen, & Lönnqvist, 2018	624	Finland	ESFPQ & †***Other	-.148	5	32	CS
Lenhart & Neyer, 2006	208	Germany	NEO-FFI & RAS	-.260	NR	25	LG
Lester, Haig, & Monello, 1989	60	USA	EPQ & MDS	-.198	11	34	CS
Luo, 2009	234	USA	BFI & †****Other	-.128	NR	19	CS
Luo & Klohnen, 2005	582	USA	PANAS, BFI, & MAT	-.135	0	28	CS
Madoures, 2018	323	Germany & UK	BFI & RAS	-.070	NR	66	CS
Mattson, Rogge, Johnson, Davidson, & Fincham, 2012	1656	USA & Canada	EPQ-N, MASQ, PN-SMD, CSI, & ISS	-.320	4	28	LG
McDaniel, Galovan, Cravens, & Drouin, 2018	358	USA	NEO-FFI & QMI	-.256	10	32	CS
McNulty, 2008	144	USA	BFI & QMI	-.273	0	24	CS
Mercado Garcia, 2013	1499	USA/Mexico	NEO-PI-R & RAS	-.111	NR	23	CS
Miller, Mason, Canlas, Wang, Nelson, & Hart, 2013	467	China	CES-D & KMSS	-.228	NR	34	CS
Möller, 2004	212	Sweden	HSPQ, EPQ-I, & †*****Other	-.342	NR	37	LG
Mosher, 2001	238	Canada	EPQ-N, DAS, & EMS	-.093	2	21	LG

Najarpourian, Fatehizadeh, Etemadi, Ghasemi, Abedi, & Bahrami, 2012	328	Iran	NEO-FFI & ENRICH	-.560	14	43	CS
Noftle & Shaver, 2006	285	USA	NEO-PI-R & PRQC	-.502	NR	20	CS
O'Meara & South, 2019	1964	USA	BFM & ††Other	-.220	11	47	LG
O'Rourke, Smith, & DeLongis, 2012	208	Canada	NEO-FFI & DAS	-.310	65	39	CS
Odilavadze, Panjikidze, Martshvishvili, Mestvirishvili, & Kvitsiani, 2019	174	Georgia	HEXACO & RDAS	.235	9	33	CS
Parker, Tambling, & Campbell, 2013	223	USA	MDI & RDAS	-.470	NR	29	LG
Radev, Bogdanović, & Anđelković, 2019	199	Serbia	NEO-PI-R & DAS	-.118	NR	44	CS
Ratcliffe, 2013	875	USA	††*Other & ††**Other	-.369	NR	35	CS
Renshaw, Blais, & Smith, 2010	602	USA	NEO-PI-R & MAT	-.350	28	54	CS
Robins, Caspi, & Moffit, 2000	720	New Zealand	MPQ & ††***Other	-.225	2	21	CS
Robins, Caspi, & Moffitt, 2002	712	New Zealand	MPQ & ††****Other	-.240	2	24	LG
Rodrigues, 2010	1723	USA	EPQ-N & CSI	-.264	8	29	LG
Rodrigues, 2010	624	USA	EPQ-N, DAS, MAT, & CSI	-.350	4	29	LG
Rogers, 1999	206	USA	NEO-PI-R & DAS	-.340	15	36	CS
Rogge, Bradbury, Hahlweg, Engle, & Thurmaier, 2006	170	Germany	BSI & MAT	-.295	4	28	LG
Rosowsky, King, Coolidge, Rhoades, & Segal, 2012	64	USA	NEO-FFI & CMSS	-.485	49	73	CS
Russell & McNulty, 2011	144	USA	BFI & QMI	-.273	0	24	LG
Russell & Wells, 1994	2400	UK	EPQ & †††Other	-.150	14	37	CS
Russell & Wells, 1994	188	Britain	EPQ-R & †††*Other	-.280	14	38	CS

Saeed Abbasi, Rattan, Kousar, & Khalifa Elsayed, 2018	819	USA	BFI & MDS	-.007	NR	27	CS
Schaffhuser, Allemand, & Martin, 2014	432	Switzerland	BFI, BFI-K, & RAS	-.107	23	48	CS
Schaffhuser, Wager, Lüdtke, & Allemand, 2014	282	Switzerland	BFI & RAS	-.180	24	50	LG
Scollon & Diener, 2006	1130	Australia	EPQ & †††**Other	-.510	NR	38	LG
Scott & Cordova, 2002	182	USA	BDI & DAS	-.435	11	39	CS
Sharma & Raju, 2013	110	India	NEO-FFI & RAS	-.280	1	25	CS
Shaver & Brennan, 1992	80	USA	NEO-PI & RRF	-.330	NR	19	LG
Slatcher & Vazire, 2009	120	USA	TIPI, BFI, & RAS	-.263	2	21	CS
Slatcher & Vazire, 2009	110	USA	TIPI & RAS	-.166	1	29	LG
Smith, Jarnecke, & South, 2020	202	USA	PID & DAS	-.082	0	27	LG
Solomon & Jackson, 2014	8206	Australia	BFI & ††††Other	-.210	23	50	LG
Sohrabi & Narimani, 2018	150	Iran	HEXACO & EMS	-.370	5	38	CS
Sousou, 2004	146	USA	NEO-FFI, MSI-R	.073	10	38	CS
Spiker, Hammer, & Parnell, 2018	206	USA	BFI-44 & CSI-4	-.140	16	46	CS
Stroud, Durban, Saigal, & Knobloch-Fedders, 2010	236	USA	MPQ & MSI-R	-.344	9	37	CS
Sullivan, 1997	344	USA	NEO-PI, EPQ-N, DIM, MAT, QMI, SD	-.105	0	27	LG
Taggart, Bannon, & Hammett, 2019	116	USA	BFI-10 & DAS-7	-.181	2	21	LG
Tandler, Krüger, & Petersen, 2021	163	Germany	BFI-25 & RAS	-.170	16	41	CS
Taormina & Ho, 2012	258	China	††††*Other & †††††Other	-.140	NR	33	CS
Tong, Jia, He, Lan, & Fang, 2021	536	China	NEO-FFI & QMI	-.231	1	29	LG
Tong, Li, Zhou, He, Ju, Li, & Fang, 2018	536	China	NEO-FFI & QMI	-.277	1	29	LG
Treviño, Wooten, & Scott, 2007	196	USA	BDI-II & DAS	-.500	11	36	CS

Wang, Kim, & Boerner, 2018	4436	USA	MDIPS, IPIP, & ⏚Other	-.128	37	67	CS
Wang, Kim, & Stokes, 2020	6356	USA	MDIPS, IPIP, & ⏚*Other	-.164	37	67	LG
Wang, Wang, Xie, Wang, Wang, Nie, & Lei, 2018	429	China	CES-D & ⏚**Other	-.320	NR	NR	CS
Wang, Xie, Wang, Wang, & Lei, 2017	243	China	CES-D & ⏚***Other	-.360	NR	36	CS
Watson, Hubbard, & Wiese, 2000	148	USA	NEO-FFI & MAT	-.371	17	47	CS
Weidenbach, 2013	136	USA	NEO-FFI & MAT	-.196	15	41	CS
Weidmann, Ledermann, & Grob, 2016	756	Switzerland	BFI & RAS	-.109	24	29	LG
Weidmann, Schönbrodt, Ledermann, & Grob, 2017	474	Switzerland	BFI & RAS	-.129	24	49	LG
Whisman, Uebelacker, & Weinstock, 2004	1488	USA	MMPI-2 (ANX & DEP) & DAS	-.241	16	41	CS
White, Hendrick, & Hendrick, 2004	196	USA	NEO-PI-R & RAS	-.327	NR	NR	CS
Whiteford, 2011	118	USA	NEO-FFI & QMI	-.186	7	33	CS
Whitton & Kuryluk, 2012	484	USA	CES-D & CSI	-.263	1	19	CS
Whitton & Kuryluk, 2014	571	USA & Puerto Rico	CES-D & CSI-4	-.330	7	41	CS
Wilson, Elkins, Bair, Oleynick, Malone, McGue, & Iacono, 2018	284	USA	PID-5 & DAS	-.400	NR	32	LG
Wodzidlo & Segrin, 2013	372	USA	BFI & QMI	-.203	2	26	CS
Wollny, Jacobs, & Pabel, 2020	272	Germany	BFI & IMS	-.167	9	34	CS
Van Scheppingen, Chopik, Vleidorn, & Denissen, 2019	8928	USA	MDIPS & ⏚****Other	-.185	37	67	LG
Vandermeer, Kotelnikova, Simms, & Hayden, 2018	752	USA	NEO-FFI & DAS	-.401	NR	34	CS
Vater & Schröder-Abé, 2015	274	Germany	BFI & RAS	-.138	10	32	CS

Verreault, Sabourin, Lussier, Normandin, & Clarkin, 2013	744	Canada	NEO-FFI & DAS	-.246	7	28	CS
Vossenkemper, 2020	368	USA	BFI, MAT, QMI, & MAS	-.300	NR	NR	CS
Yousefian Tehrani, 2015	1506	USA	RELATE	.017	NR	25	CS
Zare, Nasir, & Mastor, 2012	300	Iran	NEO-FFI & ENRICH	-.387	NR	NR	CS
Zimet, 2001	1268	USA	BFI & MAT	-.233	2	27	CS
Zoby, 2005	329	USA	BFI-44 & KMSS	-.180	4	28	CS

Note. CS = Cross-sectional; LG = Longitudinal; NR = Not Reported; BFI = Big Five Inventory; QMI = Quality Marriage Index; RAS = Relationship Assessment Scale; BDI = Beck Depression Inventory; EPQ-N = Neuroticism Scale of the Eysenck Personality Questionnaire; MAT = Marital Adjustment Test; KMSS = Kansas Marital Satisfaction Scale; SD = Semantic Differential; MDS = Marital Dissatisfaction Scale; NEO-FFI = shortened version of the NEO Personality Inventory; PANAS-X = Positive and Negative Affect Schedule – Expanded Form; ORS = Overall Relationship Satisfaction; SCL-90-R = Symptom Checklist-90-R; DAS = Dyadic Adjustment Scale; MPQ-BF = 155-item Multidimensional Personality Questionnaire – Brief Form; MMPI-2 (ANX & DEP) = Minnesota Multiphasic Personality Inventory – 2 Anxiety and Depression subscales; NEO-PI-R = NEO Personality Inventory – Revised; CES-D = Center for Epidemiological Studies-Depression Scale; EMS = Enrich Marital Satisfaction Scale; EPQ-R = Eysenck Personality Questionnaire-Revised; SMAT = Short Marital Adjustment Test; RELATE = RELATE Emotion Scale; CSI = Couples Satisfaction Index; PRS = Personality Rating Scale; HADS = Hospital Depression and Anxiety Scale; BFI-K = BFI Short Version; HSPQ = High School Personality Questionnaire; MDRCI = Multidimensional Determinants of Relationship Commitment Inventory; SLS = Satisfaction With Life Scale; CRSI = Conflict Resolution Styles Inventory; BSI = Hostility, Interpersonal Sensitivity, Depression, and Anxiety scales of the Brief Symptoms Inventory; BABS = Bradburn Affect Balance Scale; AAS = Adult Attachment

Scale; ERS = Ego-Resiliency Scale; LWMAT = Locke-Wallace Marital Adjustment Test, PSSI = Pinney Sexual Satisfactory Inventory; MSI-R = Marital Satisfaction Inventory – Revised; IPIP = International Personality Item Pool Scales of Neuroticism, Agreeableness, and Conscientiousness; CMSS = Comprehensive Marital Satisfaction Scale; BFM = Big Five Model traits; ESFPQ = Extra Short Five Personality Questionnaire; GTS = General Temperament Survey; RQI = Relationship Quality Interview; FCFFMQ = Forced Choice Five Factor Markers Questionnaire; MDIPS = Midlife Development Inventory Personality Scales; MMPI-2-RF = Minnesota Multiphasic Personality Inventory – 2 – Restructured Form; IMS = Investment Model Scale; HEXACO = HEXACO Personality Inventory; RDAS = Revised Dyadic Adjustment Scale; PID-5 = Personality Inventory for DSM-5; IPIP-NEO = Neuroticism sub-scale of the IPIP; MAS = Marital Attitude Survey; 3M-40 = Big Five; PD-I = Personality Domain Inventory; QPD = Questionnaire for Partnership Diagnostics; MASQ = Mood and Anxiety Symptom Questionnaire; BAI = Beck Anxiety Inventory; CA = Corumination Questionnaire; QRI = Quality of Relationships Inventory; DPQ = Dutch Personality Questionnaire; DRQ = Dutch Relationship Questionnaire; PRQC = Perceived Relationship Quality Components Inventory; PWI = Personal Well-Being Index; MDI = Major Depression Inventory; RRF = Relationship Rating Form; TIPI = Ten-Item Personality Inventory; 16PF = 16 Personality Factor Questionnaire; MOQ = Marital Opinion Questionnaire; IPQ = Iowa Personality Questionnaire; MSS = Marital Satisfaction Scale; DIM = Dysfunctional Impulsivity Measure; GMQ = Good Marriage Questionnaire; SWLS = Satisfaction With Life Scale; PN-SMD = Positive and negative relationship satisfaction; ISS = Index of Sexual Satisfaction; *Other = Four items from the University of Southern California Longitudinal Study of Three-Generation Families measures of positive affect; **Other = A measure of trait neuroticism; ***Other = "How satisfied are you with your marriage or romantic relationship?"; ****Other = Relationship quality was

measured using a 4-item scale that was based on items developed for the Iowa Youth and Families Project; *****Other = "Relationship satisfaction was assessed via two questions, one asking how “happy” (1 – extremely unhappy to 7 - extremely happy) and one asking how “satisfied” (1 – completely dissatisfied to 7 - completely satisfied) respondents were with their current relationship (Surjadi et al., 2011)"; ¦Other = Marital quality was measured using 2 items that tapped the spouses’ overall feelings of happiness and satisfaction with the marriage. Spouses were asked to respond to how happy they were with their marital relationship using a 6-point scale (0, extremely unhappy and 5, extremely happy) and to respond to how dissatisfied they were with their relationship using a 5-point scale (1, completely satisfied and 5, not at all satisfied); ¦*Other = Sample 1: Marital satisfaction was assessed in the BHPS by asking respondents to report how satisfied or dissatisfied they felt about their husband or wife using a seven point scale from 1 ―not at all satisfied‖ to 7 ―completely satisfied. Sample 2: Marital satisfaction was assessed by asking respondents to report how satisfied or dissatisfied they felt about their relationship with their partner using an eleven-point scale (0 = ―completely dissatisfied‖ to 10 = ―completely satisfied, M = 8.39, SD = 1.84); ¦**Other = Marital satisfaction was measured by a single item with a 7-point scale running from extraordinarily happy to extremely unhappy; ¦***Other = One item on a scale of 1-10: How satisfied are you in your relationship with your spouse?; ¦****Other = Participants indicated their overall satisfaction on a single item: “all things considered, how happy are you in your relationship?” using a 10-point scale ranging from “very unhappy” to “perfectly happy”. They also completed a 10-item relationship satisfaction scale adapted from Hendrick’s (1988) Relationship Assessment Scale using a 7-point scale (ranging from strongly disagree to strongly agree); ¦*****Other = The individual items making up the broad partner-satisfaction measure were based on six open-ended questions: “How would you characterize your spouse

(1 = only negative characteristics men- tioned, 5 = only positive characteristics mentioned)?", "Do you and your spouse have any particular interests that hold you together (1 = no common leisure-time interests, no recreation together, 5 = share each other's interests, same amusements)?", "To give an overall impression of your relationship, how would you describe the home atmosphere (1 = very disharmonious, almost divorce atmosphere, 6 = unusually cordial relations, attitudes in harmony, open and warm home atmosphere)?", "How often do you spontaneously cuddle or caress each other (1 = seldom, 5 = daily)?", "Are you sexually well adjusted to each other (1 = have no sexual life together, or very Table 1. Means and standard deviations of neuroticism scores Females seldom, 6 = very well adjusted to each other)?", and "Do you receive encouragement and support from your partner when you have prob- lems at work (1 = partner is more of an obstacle than a source of support, 5 = receive all the help I need)?". A brief questionnaire about partner relations contained four main questions: "Does your partner talk to you about his/her prob- lems (1 = never, 4 = always)?", "How warm are your feelings for your partner (1 = no warm feelings at all, 5 = very warm feelings)?", "How do you and your partner get along together (1 = badly, 5 = very well)?", and "How often do you get really angry with your partner (1 = very often, to 5 = very seldom)?" ††Other = "For the current study, we used a scale (South & Krueger, 2008) composed of 21 items measuring relationship disagreement , spousal support/strain, relationship risk, and relationship decision making"; ††*Other = 7-item scale, asking participants "How much do these words or phrases describe you?" Words or phrases included "sad and blue", "feel hopeless", "depressed", "fearful", "tense", "nervous", and "worrier"; ††**Other = Marital satisfaction was measured with a 7-item scale, assessing the degree of satisfaction individuals felt in different aspects of their marriage; ††***Other = Relationship Satisfaction was assessed using 14 interview questions that asked how satisfied the respondent was with different domains of the relationship;

††****Other = Relationship quality was assessed using 28 interview questions that ask about shared activities and interests, the balance of power, respect and fairness, emotional intimacy and trust, and open communication; †††Other = Six questions concerning feelings of closeness to one's partner, enjoyment of their company and so on, with five-point ratings for the response categories; †††*Other = "Are you happy?" and rate the happiness of your spouse; †††**Other = Five items assessed satisfaction with one's romantic relationship; ††††Other = “We assessed relationship satisfac- tion at each wave (Waves 1–5) using a one-item question that asked participants how satisfied they were with their romantic partner on a 0 –10 scale, in which 10 denoted Completely satisfied (Dyrenforth et al., 2010)”; ††††*Other = This was measured with a five-item scale that focused on the prominent anxiety/worry facet of neuroticism. Two items were from Costa and McCrae's (1992) Neuroticism domain; for example, ‘I am filled with doubts about things’ (alpha = .91). The remaining three items were from Peterson and Seligman's (2004) Neuroticism measure; for example, ‘I usually expect the worst’; †††††Other = RS- Emotional: two of the five items were adapted from Sternberg's (1997) 12-item Triangular Love Scale (TLS), one item adapted from Holt Relationship Intimacy Questionnaire, one item from Levinger, Rands, and Talaber involvement scale, one newly created. RS-Intellectual: Three of the five items for this scale were adapted from Levinger et al.'s (1977) involvement scale, namely, ‘My partner and I enjoy sharing ideas with each other’, ‘My partner and I understand each other very well’, and ‘My partner and I have countless things to talk about’ (no reliability reported). The remaining two items were newly created RS- Spiritual: all five newly created items; RS- Physical: four selected from Holt's HRIQ, one newly created; Belongingness Need was measured by a newly created 12-item scale developed from Maslow's (1943, 1971) theory using statements that focused on satisfaction with intimacy, support, warmth, and affection that one has with a spouse/partner, family, friends, and

associates; ⏚Other = Marital quality measured as 7 items on spousal support and spousal strain; ⏚*Other = Two-dimensional scale measuring spousal support and spousal strain; ⏚**Other = The 4-item satisfaction scale was used to test adults' RS, which was developed by Murray, Holmes, Griffin, and Derrick (2015); ⏚***Other = The four-item satisfaction scale indexed participants' global evaluation of their relationships quality (Murray, Holmes, Griffin, & Derrick, 2015); ⏚****Other = three items: “How much do they (i.e., the spouse) really understand the way you feel about things?,” “How much can you rely on them if you have a serious problem?,” and “How much can you open up to them if you need to talk about your worries?”