

Codevelopment of Well-Being and Self-Esteem in Romantic Partners:
Disentangling the Effects of Mutual Influence and Shared Environment

Ulrich Orth and Ruth Yasemin Erol

University of Bern

Thomas Ledermann

Utah State University

Alexander Grob

University of Basel

© American Psychological Association. This article has been accepted for publication but has not been through the copyediting, typesetting, pagination, and proofreading process. This article may not exactly replicate the final, authoritative version published in the journal. It is not the copy of record. Please cite this article as follows:

Orth, U., Erol, R. Y., Ledermann, T., & Grob, A. (2018). Codevelopment of well-being and self-esteem in romantic partners: Disentangling the effects of mutual influence and shared environment. *Developmental Psychology*, *54*, 151-166.
<http://dx.doi.org/10.1037/dev0000400>

Author Note

Ulrich Orth and Ruth Yasemin Erol, Department of Psychology, University of Bern; Thomas Ledermann, Department of Psychology, Utah State University; Alexander Grob, Department of Psychology, University of Basel.

This research was supported by Swiss National Science Foundation Grants PP00P1-123370, CRSII1-130432, and CRSII1-147614. Moreover, this research used the Longitudinal Study of Dual-Earner Couples, 1989–1992 data set (made accessible in 1993 in original paper record and numeric file form). These data were collected by Rosalind C. Barnett, and are available through the Henry A. Murray Research Archive of the Institute for Quantitative Social Science at Harvard University, Cambridge, Massachusetts [Producer and Distributor].

Correspondence concerning this article should be addressed to Ulrich Orth, Department of Psychology, University of Bern, Fabrikstrasse 8, 3012 Bern, Switzerland. E-mail: ulrich.orth@psy.unibe.ch

Abstract

Common sense suggests that romantic partners tend to be interdependent in their well-being and self-esteem. We tested the degree to which codevelopment in romantic partners (i.e., development in similar directions) is due to mutual influence between partners or due to the effects of shared environment, using longitudinal data from five samples of couples ($N = 4,116$ participants). The samples included dating, cohabiting, and married couples, with a broad age range covering young adulthood to old age. We used a longitudinal version of the actor-partner interdependence model and examined measures of well-being (life satisfaction, positive affect, negative affect, and depression) and self-esteem. After conducting the study-level analyses, we meta-analytically aggregated the findings across studies. The results showed significant mutual influence between partners in life satisfaction, positive affect, negative affect, and depression, and significant effects of shared environment on life satisfaction, positive affect, negative affect, depression, and self-esteem. The findings suggest that both mutual influence and shared environment account for codevelopment of well-being in romantic partners. In contrast, only shared environment but not mutual influence contributes to the partners' codevelopment of self-esteem. The findings have important implications because they show that the level of well-being of a person's relationship partner has long-term influence on the person's own well-being.

Keywords: close relationships, well-being, self-esteem, longitudinal, actor-partner interdependence model

Codevelopment of Well-Being and Self-Esteem in Romantic Partners:
Disentangling the Effects of Mutual Influence and Shared Environment

Many psychologists and laypeople assume that the well-being of two people in a romantic relationship is highly interdependent. Moreover, cultural rites such as the Christian marriage vow (“for better, for worse”) emphasize that relationship partners are expected to go together through good times and bad times, as well as through sickness and health. Thus, codevelopment (i.e., development in similar directions) of the well-being of relationship partners even has a normative component. However, although research clearly shows that relationship partners have similar levels of well-being (for a review, see Meyler, Stimpson, & Peek, 2007) and although some studies suggest that partners emotionally converge over time (Anderson, Keltner, & John, 2003; but see Gonzaga, Campos, & Bradbury, 2007), the available evidence does not tell us to what degree codevelopment of romantic partners is due to mutual influence between partners or to the effects of shared environmental factors.

In the present research, we therefore test for the effects of both mutual influence and shared environment, and disentangle these two sources of codevelopment, using longitudinal data from both partners of couples. Well-being was assessed using measures of life satisfaction, positive affect, and negative affect (following the definition of well-being by E. Diener, 1984, 2000) and using measures of depression, which is a key indicator of negative affect and particularly important given its clinical relevance (Berman, 2009; Kessler et al., 2005). Besides measures of well-being, we also examined self-esteem. Although self-esteem is not considered an indicator of well-being in the narrower sense (Busseri & Sadava, 2011; E. Diener, 1984; E. Diener, Suh, Lucas, & Smith, 1999), it is closely related to life satisfaction, positive affect,

negative affect, and depression (Lucas, Diener, & Suh, 1996; Orth, Robins, & Widaman, 2012; Robins, Hendin, & Trzesniewski, 2001; Sowislo & Orth, 2013).

Identifying the factors that influence the development of well-being and self-esteem in adulthood is important. Research suggests that people's habitual level of well-being shows gradual changes across the life span (Carstensen et al., 2011; Mroczek & Spiro, 2005). For example, negative affectivity decreases from young to middle adulthood and is relatively low during old age (Charles, Reynolds, & Gatz, 2001; Mroczek & Kolarz, 1998). Similarly, self-esteem shows normative change across adulthood, increasing from young adulthood to midlife and decreasing in old age (Orth, Maes, & Schmitt, 2015; Orth, Trzesniewski, & Robins, 2010). However, these studies also indicate that people differ significantly in the individual trajectory they follow. Research shows that romantic relationships are a central life domain in adulthood, accounting for individual differences in the development of well-being and self-esteem (M. L. Diener & Diener McGavran, 2008; Luciano & Orth, 2017; Wagner, Becker, Lüdtkke, & Trautwein, 2015). Therefore, it is important to understand how the development of these characteristics is linked to those of the relationship partner.

Theoretical Perspectives on Codevelopment in Well-Being and Self-Esteem

As mentioned above, a first explanation for why romantic partners develop in similar ways while in a relationship with each other is that there is mutual influence between the two partners. A prominent theory that makes this assumption is emotional contagion theory (Hatfield, Cacioppo, & Rapson, 1993). This theory proposes that when people interact with each other, they automatically mimic and synchronize expressions and movements with those of the interaction partner and, consequently, converge emotionally (Hatfield et al., 1993). Thus, the theory suggests that emotions are transferred from one person to the other by empathic

processes, which implies interdependence between the two persons—that is, the emotional state of one person influences the emotional state of the other person. If the contagion process occurs repeatedly because the two persons interact on a regular basis (e.g., because they are in a romantic relationship with each other), the theory suggests that contagion occurs not only with regard to emotional states but also emotional traits, such as positive and negative affectivity. Similar to emotional contagion theory, Coyne's (1976) interactional theory of depression proposes that depression might be contagious among relationship partners. Katz, Beach, and Joiner (1999) suggest that depression contagion is particularly strong when depressed individuals show excessive reassurance seeking, thereby putting relationship partners at risk for heightened negative affect and depression. Likewise, Kelley and Thibaut's (1978; see also Rusbult & Buunk, 1993; Rusbult & Van Lange, 2003) interdependence theory posits that relationship partners mutually influence each other. This theory states that the outcomes of partners become intertwined over time, that is, a person's outcome is not only the result of his or her own action but also of his or her partner's action. For example, a person's well-being is enhanced not only by his or her own successes but also by those of the partner. Taken together, emotional contagion theory, interactional theory of depression, and interdependence theory lead to the hypothesis that there is mutual influence between partners with regard to their well-being and, potentially, self-esteem.

A second explanation of codevelopment is that relationship partners are influenced by shared environment (e.g., Caspi, Herbener, & Ozer, 1992; Schimmack & Lucas, 2010; Tower & Kasl, 1996). For example, relationship partners typically share environmental factors such as family conditions (e.g., characteristics of children, contact with relatives), residence and residential area (e.g., living in a noisy and unsafe area), financial conditions (e.g., family income

and assets), stressful life events (e.g., loss of a child, unemployment, and burglary), and quality and size of social network. Although these environmental factors do not necessarily influence both partners in identical ways, this perspective suggests that shared environment leads to similarity rather than dissimilarity in both well-being and self-esteem between partners over time. Importantly, however, this perspective offers an explanation as to why partners develop in similar ways without assuming that one partner causally influences the other partner and vice versa (i.e., mutual influence).

Evidence on Codevelopment in Well-Being and Self-Esteem

In the following sections, we review the available evidence on codevelopment in relationship partners with regard to well-being (i.e., life satisfaction, positive affect, negative affect, and depression) and self-esteem.

Life Satisfaction

Cross-sectional evidence suggests that relationship partners are similar in their level of life satisfaction. For example, Bookwala and Schulz (1996) reported a medium-sized positive correlation between the life satisfaction scores of partners. Similarly, in a study with a large Indian sample, life satisfaction was positively correlated between partners with about medium effect size (Shakya, 2015). The evidence from further cross-sectional studies supports that relationship partners have similar levels of life satisfaction (Demerouti, Bakker, & Schaufeli, 2005; Peek, Stimpson, Townsend, & Markides, 2006). Moreover, Tambs and Moum (1992) tested whether the correlation of life satisfaction between partners depends on the length of marriage and found that it did not. Given that codevelopment in romantic partners would lead to increasing levels of similarity, the findings of Tambs and Moum (1992) do not support the codevelopment hypothesis for life satisfaction. However, because of the cross-sectional design of

the study, cohort effects could have confounded the results. For example, it is possible that codevelopment exists, leading to increasing similarity over time, but that couples from older cohorts had been less similar at the time of marriage due to cohort differences in assortative mating (i.e., the degree to which individuals choose relationship partners who are similar to themselves; Luo & Klohnen, 2005; Watson et al., 2004).

The association between life satisfaction of relationship partners has also been examined in longitudinal studies. For example, Schimmack and Lucas (2010) used data collected over 22 years and found evidence that life satisfaction of relationship partners changed in the same direction. Likewise, Hoppmann, Gerstorf, Willis, and Schaie (2011) examined the development of happiness—a measure conceptually related to life satisfaction—among married partners. The results showed that partners were similar with regard to both level and slope of the happiness trajectory. Also, in a longitudinal study with a nationally representative sample of couples from Germany, the life satisfaction of partners was correlated both before and after becoming unemployed (Luhmann, Weiss, Hosoya, & Eid, 2014). It is important to note that, although these studies were longitudinal, they do not inform about the potential causes of codevelopment in romantic relationships. Put differently, even though there is evidence of correlated change in life satisfaction within romantic relationships, it is unclear whether this association results from mutual influence between partners or shared environmental factors.

Positive and Negative Affect

Cross-sectional evidence indicates that romantic partners are similar in their positive and negative affectivity, at about medium effect size (Gaunt, 2006; Murdock, Lovejoy, & Oddi, 2014; Segrin et al., 2005). Moreover, in a series of longitudinal studies, Segrin and colleagues tested for prospective partner effects (i.e., mutual influence).¹ In a study with 153 couples, Segrin

et al. (2005) found that a person's negative affect did not significantly predict the partner's level of negative affect. Segrin, Badger, Dorros, Meek, and Lopez (2007) tested for partner (i.e., between-person) effects in a sample of 96 women with breast cancer and their partners; however, only one out of four partner effects was statistically significant. Segrin, Badger, and Harrington (2012) examined a sample of 70 men with prostate cancer and their partners; again, most of the partner effects tested were nonsignificant. In sum, evidence from these longitudinal studies does not consistently support the notion that there is mutual influence between relationship partners with regard to positive and negative affect. However, it should be noted that, although these studies provide relevant information on the association of affective measures between partners, the sample sizes might not have provided sufficient power to reliably test for partner effects.

In addition, some studies examined diary data collected over short periods. For example, in a study with 30 married couples, the mood of partners was assessed four times per day across three days (Saxbe & Repetti, 2010). Whereas negative mood was associated with the partner's negative mood, positive mood was not associated between partners, raising the possibility that interdependence between partners might be stronger for negative affect than for positive affect. However, this diary study tested for concurrent and not prospective associations between partners, which does not provide evidence with regard to influences between partners. Song, Foo, and Uy (2008) assessed prospective partner effects in 50 dual-earner couples and found significant partner effects for both positive and negative mood; however, the effects were larger when the partners were physically together and when the time interval between the partners' reports was short. To summarize, the available evidence does not allow for clear-cut conclusions about the factors that account for long-term codevelopment of positive and negative affect.

Depression

Cross-sectional evidence clearly shows that depression is correlated between relationship partners (e.g., Du Rocher Schudlich, Papp, & Cummings, 2011; Joiner & Katz, 1999; Townsend, Miller, & Guo, 2001; Whiffen & Aube, 1999; but see Segrin, Powell, Givertz, & Brackin, 2003). Several longitudinal studies have tested for prospective effects of depression between partners. For example, Tower and Kasl (1996) examined a sample of older married couples and found that change in depressive symptoms in one partner had an effect on the other partner's depressive symptoms three years later. Similarly, using a large sample of older married couples, Siegel, Bradley, Gallo, and Kasl (2004) found evidence for significant partner effects in depression. There were also some gender differences. Whereas Gerstorf, Hoppmann, Kadlec, and McArdle (2009) found a prospective effect from the female to the male partner, Kouros and Cummings (2010) found an effect from the male to the female partner. In sum, although the evidence suggests that relationship partners might be interdependent in their level of depression, it is unclear whether the effect holds across gender or whether it is smaller, or even absent, for one of the genders.

Self-Esteem

Cross-sectional research indicates that there is a small to medium-sized correlation between partners' self-esteem levels (e.g., Erol & Orth, 2013; Murray, Holmes, & Griffin, 2000; R. B. Schafer & Keith, 1992; Schumm, Figley, & Fuhs, 1980; Shackelford, 2001). Moreover, longitudinal research has tested whether romantic partners show correlated change in self-esteem, using data from two large samples of couples across 12 and 15 years, respectively (Erol & Orth, 2014). Only one of the samples showed evidence of correlated change within couples. However, little evidence is available regarding the factors that account for codevelopment of self-esteem in romantic relationships. Wagner, Voelkle, Hoppmann, Luszcz, and Gerstorf (2017)

found that the male partner's self-esteem predicted the female partner's self-esteem one year later but not vice versa. Moreover, Schaffhuser, Wagner, Lüdtkke, and Allemann (2014) tested for mutual influence in self-esteem but did not find evidence of significant effects.

The Present Research

The aim of the present research was to examine the factors that account for codevelopment in romantic partners with regard to their well-being and self-esteem. More precisely, we tested for the effects of two fundamentally different sources of codevelopment, that is, mutual influence between partners and shared environment. Whereas one perspective (i.e., mutual influence) is based on the assumption that partners causally influence each other, the other perspective (i.e., shared environment) does not require assumptions about causal influence between partners. Although the two perspectives make divergent predictions with regard to the cause of codevelopment, they are not necessarily mutually exclusive. In fact, it is possible that both processes (i.e., mutual influence between partners and effects of shared environment) operate simultaneously. Drawing from the theoretical perspectives reviewed above (such as emotional contagion theory and interdependence theory; Hatfield et al., 1993; Kelley & Thibaut, 1978), we hypothesized that there is mutual influence in well-being and self-esteem. Moreover, based on prior research on shared environmental effects (e.g., Caspi et al., 1992; Schimmack & Lucas, 2010), we hypothesized that shared environment accounts for codevelopment in well-being and self-esteem, over and above the effects of mutual influence.

For the analyses, we used data from five studies with samples of couples to gain more robust estimates of the effects. After computing the study-level estimates, we meta-analytically aggregated the findings across studies. Such cross-study aggregation helps by reducing biases (which may result from peculiarities within single studies), increases the power of statistical

tests, and increases the precision of estimates, as indicated by smaller confidence intervals (Goh, Hall, & Rosenthal, 2016). The samples were relatively diverse and came from two different countries (i.e., the United States and Switzerland), strengthening the generalizability of the results.

Method

The five longitudinal studies used in this research included data from both partners of couples.² For couples to be included in the present analyses, each partner had to provide data on at least one of the constructs examined in this research. Table 1 provides an overview of descriptive information on the studies. Table 2 provides the means and standard deviations of the measures used in the present research separately for male and female partners (for intercorrelations among measures, see the supplemental materials, Tables S1 to S5).

Study 1: Longitudinal Study of Generations (LSG)

Participants. The LSG includes members of families that were randomly drawn from a subscriber list of a large health maintenance organization in Southern California (Bengtson, 2009). Participants were assessed in 1971, 1985, 1988, 1991, 1994, 1997, and 2000. However, because the interval between 1971 and 1985 was relatively long (14 years), whereas the intervals between the other waves were shorter and of equal length (3 years), and because in 1971 the LSG did not include the depression measure and the full self-esteem measure, in the present research we used data of the six waves from 1985 to 2000 (in this article denoted as Waves 1 to 6). The present sample included all couples who provided data at least at one of the six waves from 1985 to 2000. The sample consisted of 885 couples ($N = 1,770$ participants), including married and cohabiting couples. At Wave 1, the mean length of relationship was 23.3 years (SD

= 15.9; range 1–62), the mean age of men was 53.9 years ($SD = 15.7$; range 27–88), and the mean age of women was 51.7 years ($SD = 15.9$; range 22–93).

Measures.

Positive and negative affect. Positive and negative affect were assessed with the Affect Balance Scale (Bradburn, 1969), with five items measuring positive affect and five items measuring negative affect. The validity of the scale has been repeatedly supported (Harding, 1982; Kim & Mueller, 2001). For each item, participants reported whether they ever felt that way during the past week (0 = *no*; 1 = *yes*), with $M = 0.77$ ($SD = 0.26$) for positive affect and $M = 0.29$ ($SD = 0.29$) for negative affect averaged across waves. Coefficient alpha can underestimate the reliability of scales when items are dichotomous. We therefore used the method proposed by Raykov, Dimitrov, and Asparouhov (2010) for estimating the reliability of scales with dichotomous items. For positive affect, the reliability ranged from .78 to .83 (male partners) and from .81 to .85 (female partners) across waves. For negative affect, the reliability ranged from .64 to .76 (male partners) and from .65 to .72 (female partners) across waves.

Depression. Depression was assessed with the 20-item Center for Epidemiologic Studies Depression Scale (CES-D; Radloff, 1977). The CES-D is a frequently used self-report measure for the assessment of depressive symptoms in non-clinical, sub-clinical, and clinical populations, and its validity has been repeatedly confirmed (Eaton, Smith, Ybarra, Muntaner, & Tien, 2004). Participants reported how frequently they experienced depressive symptoms during the past week. Responses were measured on a 4-point scale (0 = *rarely or none of the time*; 3 = *most or all of the time*), with $M = 1.48$ ($SD = 0.41$) averaged across waves. The alpha reliability ranged from .86 to .89 (male partners) and from .86 to .90 (female partners) across waves.

Self-esteem. Self-esteem was assessed with the 10-item Rosenberg Self-Esteem Scale (Rosenberg, 1965), a commonly used and well-validated measure of self-esteem (Robins et al., 2001). Responses were measured using a 4-point scale (1 = *strongly disagree*; 4 = *strongly agree*), with $M = 3.42$ ($SD = 0.43$) averaged across waves. The alpha reliability ranged from .82 to .87 (male partners) and from .81 to .86 (female partners) across waves. In the LSG, the 10-item RSE was used only from 1988 to 2000, but not in 1985; therefore, the analyses for self-esteem were based on five waves (however, for all other measures of the LSG, the analyses were based on six waves).

Attrition analyses. Data on study variables were available for 635 couples at Wave 1, 692 couples at Wave 2, 706 couples at Wave 3, 758 couples at Wave 4, 719 couples at Wave 5, and 710 couples at Wave 6.³ To examine the potential impact of attrition, we compared individuals who participated in the last wave (i.e., Wave 6) with individuals who were part of the sample but did not participate in the last wave (thereby operationalizing the group of participants who dropped out of the study). These two groups were compared on study variables assessed at the first wave (in the LSG, an exception is that for self-esteem we used the Wave 2 data for the attrition analyses; as noted above, this construct was not assessed at Wave 1). In this study and all other studies including more than two waves of data, we tested for attrition effects only with regard to the last wave. Our rationale for this procedure was as follows: If there is a lack of attrition effects with regard to the group of participants who dropped out anytime between the first and last wave (e.g., between Wave 1 and Wave 6 of the LSG), then it is likely that there is a similar lack of attrition effects for the group of participants who dropped out before earlier waves (e.g., between Wave 1 and Wave 3 of the LSG; thus, the latter group of participants is a subgroup of the former group of participants). This reasoning suggests that it is unnecessary to

test for attrition effects with regard to intermediate waves. Apart from that, we argue that the potential impact of attrition effects is significantly reduced in this research given that in the structural equation modeling analyses we used full information maximum likelihood (FIML) for dealing with missing data (J. L. Schafer & Graham, 2002; Widaman, 2006). In research situations with longitudinal dropout, J. L. Schafer and Graham (2002) recommend using all available data and FIML because a significant amount of missing information can then be recovered from those waves at which information on participants is available. Moreover, even if data are not missing completely at random, FIML produces less biased, more reliable, and more valid results compared to simpler methods of dealing with missing data, such as listwise deletion or mean imputation (Allison, 2003; J. L. Schafer & Graham, 2002). In many research situations, FIML yields parameter estimates that are very close to those estimates that would have resulted if no longitudinal dropout had occurred (Widaman, 2006; Wothke, 2000). In the LSG, male participants who dropped out reported lower positive affect ($M = 0.58$ vs. 0.76 ; $d = -0.60$), higher depression ($M = 1.53$ vs. 1.40 ; $d = 0.33$), and lower self-esteem ($M = 3.36$ vs. 3.56 ; $d = -0.46$); differences in negative affect among male participants and differences in positive affect, negative affect, depression, and self-esteem among female participants were nonsignificant. Given that the majority of differences were nonsignificant and given that we used FIML for dealing with missing data, we concluded that attrition was not a critical issue in the LSG.

Study 2: Iowa Youth and Families Project (IYFP)

Participants. The IYFP includes a sample of families from rural Iowa (Conger et al., 2011). Participants were assessed at four waves in 1989, 1990, 1991, and 1992. The sample consisted of 451 couples ($N = 902$ participants), including married and cohabiting couples. At Wave 1, the mean length of relationship was 17.8 years ($SD = 3.1$; range 12–35), the mean age of

men was 39.7 years ($SD = 4.9$; range 31–68), and the mean age of women was 37.7 years ($SD = 4.1$; range 29–53).

Measures.

Positive affect. Positive affect was assessed with 6 items (Conger et al., 2011, selected the items from the General Positive Affect Scale of the Mental Health Inventory by Veit & Ware, 1983). Participants reported how they viewed their life during the past month. The items were “Have you generally enjoyed the things you do?”, “Have you felt that the future looks hopeful and promising?”, “Has your daily life been full of things that were interesting to you?”, “Did you feel relaxed and free of tension?”, “Were you a happy person?”, and “Has living been a wonderful adventure for you?” (Conger et al., 2011). Responses were measured on a 6-point scale (1 = *none of the time*; 6 = *all of the time*), with $M = 4.48$ ($SD = 0.78$) averaged across waves. The alpha reliability ranged from .84 to .90 (male partners) and from .87 to .91 (female partners) across waves.

Depression. Depression was assessed with 13 items of the Symptom Checklist-90-Revised (SCL-90-R; Derogatis & Savitz, 1999). Participants reported how frequently they experienced depressive symptoms during the past week. Responses were measured on a 5-point scale (1 = *not at all*; 5 = *extremely*), with $M = 1.43$ ($SD = 0.47$) averaged across waves. The alpha reliability ranged from .87 to .92 (for both male and female partners) across waves.

Self-esteem. Self-esteem was assessed with the 10-item RSE (Rosenberg, 1965). Responses were measured on a 5-point scale (1 = *strongly disagree*; 5 = *strongly agree*), with $M = 3.99$ ($SD = 0.53$) averaged across waves. The alpha reliability ranged from .85 to .88 (male partners) and from .88 to .90 (female partners) across waves.

Attrition analyses. Data on study variables were available for 451 couples at Wave 1, 424 couples at Wave 2, 407 couples at Wave 3, and 404 couples at Wave 4. To examine the potential impact of attrition, we compared individuals who did versus did not participate in the most recent wave (i.e., Wave 4) on study variables assessed at the first wave. No significant differences emerged. Thus, attrition was not a serious issue in the IYFP.

Study 3: Longitudinal Study of Dual-Earner Couples (LSDEC)

Participants. The LSDEC includes a sample of dual-earner couples from the greater Boston area (Barnett, 1993). Participants were assessed at three waves, at 1-year intervals starting in 1989. The sample consisted of 300 couples ($N = 600$ participants), including married and cohabiting couples. At Wave 1, the mean length of relationship was 8.3 years ($SD = 5.2$; range 0–22; in the LSDEC, “less than one year” was coded as 0), the mean age of men was 35.2 years ($SD = 4.3$; range 25–42), and the mean age of women was 34.3 years ($SD = 4.8$; range 22–49).

Measures.

Depression. Depression was assessed with 13 items of the SCL-90-R (Derogatis & Savitz, 1999). Participants reported how frequently they experienced depressive symptoms during the past week. Responses were measured on a 5-point scale (0 = *not at all*; 4 = *extremely*), with $M = 0.67$ ($SD = 0.55$) averaged across waves. The alpha reliability ranged from .86 to .90 (male partners) and from .85 to .90 (female partners) across waves.

Self-esteem. Self-esteem was assessed with the 10-item RSE (Rosenberg, 1965). Responses were measured on a 4-point scale (1 = *strongly disagree*; 4 = *strongly agree*), with $M = 3.43$ ($SD = 0.47$) averaged across waves. The alpha reliability ranged from .81 to .85 (male partners) and from .84 to .87 (female partners) across waves.

Attrition analyses. Data on study variables were available for 300 couples at Wave 1, 294 couples at Wave 2, and 278 couples at Wave 3. To examine the potential impact of attrition, we compared individuals who did versus did not participate in the most recent wave (i.e., Wave 3) on study variables assessed at the first wave. No significant differences emerged. Thus, attrition was not a serious issue in the LSDEC.

Study 4: Co-Development in Personality (CoDiP)

Participants. The CoDiP is a German-language study with a sample of couples living in urban and suburban regions of Switzerland (Furler, Gomez, & Grob, 2014). Participants were assessed in 2010 and 2012. The sample consisted of 236 couples ($N = 472$ participants), including married, cohabiting, and dating couples. At Wave 1, the mean length of relationship was 23.7 years ($SD = 17.6$; range 0.8–63), the mean age of men was 50.8 years ($SD = 20.0$; range 18–92), and the mean age of women was 48.5 years ($SD = 19.6$; range 16–86).

Measures.

Life satisfaction. Life satisfaction was assessed with the 5-item Satisfaction with Life Scale (SWLS; E. Diener, Emmons, Larsen, & Griffin, 1985). Example items were “I am satisfied with my life” or “So far I have gotten the important things I want in life.” Responses were measured on a 5-point scale (1 = *strongly disagree*; 5 = *strongly agree*), with $M = 4.04$ ($SD = 0.65$) averaged across waves. The alpha reliability ranged from .81 to .86 (male partners) and from .82 to .88 (female partners) across waves.

Positive and negative affect. Positive and negative affect were assessed with the 20-item Positive and Negative Affect Schedule (PANAS; Watson, Clark, & Tellegen, 1988). Participants reported how they felt “in general.” Responses were measured on a 5-point scale (1 = *very slightly or not at all*; 5 = *extremely*), with $M = 3.62$ ($SD = 0.54$) for positive affect and $M = 1.66$

($SD = 0.52$) for negative affect averaged across waves. For positive affect, the alpha reliability was .86 (male partners) at each wave and ranged from .83 to .87 (female partners) across waves. For negative affect, the alpha reliability ranged from .84 to .86 (male partners) and from .85 to .86 (female partners) across waves.

Self-esteem. Self-esteem was assessed with the 10-item RSE (Rosenberg, 1965). Responses were measured on a 4-point scale (1 = *strongly disagree*; 4 = *strongly agree*), with $M = 3.35$ ($SD = 0.47$) averaged across waves. The alpha reliability ranged from .80 to .83 (male partners) across waves and was .87 (female partners) at each wave.

Attrition analyses. Data on study variables were available for the full sample of couples at both Waves 1 and 2. Thus, there was no attrition in the CoDiP.

Study 5: My Partner and I (MPI)

Participants. The MPI is a German-language study of personality and well-being including a sample of couples living in Switzerland (Orth, 2013). Participants were assessed at two waves in 2011 with a six-month interval. The sample consisted of 186 couples ($N = 372$ participants), including married, cohabiting, and dating couples. At Wave 1, the mean length of relationship was 5.2 years ($SD = 5.4$; range 0.8–33), the mean age of men was 30.4 years ($SD = 9.3$; range 19–61), and the mean age of women was 27.7 years ($SD = 8.2$; range 18–59).

Measures.

Life satisfaction. Life satisfaction was assessed with the 5-item SWLS (E. Diener et al., 1985). Responses were measured on a 7-point scale (1 = *strongly disagree*; 7 = *strongly agree*), with $M = 5.06$ ($SD = 1.27$) averaged across waves. The alpha reliability ranged from .88 to .89 (male partners) and from .89 to .90 (female partners) across waves.

Depression. Depression was assessed with the 20-item CES-D (Radloff, 1977).

Participants reported how frequently they experienced depressive symptoms during the past week. Responses were measured on a 4-point scale (0 = *rarely or none of the time*; 3 = *most or all of the time*), with $M = 0.58$ ($SD = 0.43$) averaged across waves. The alpha reliability ranged from .86 to .88 (male partners) and from .89 to .91 (female partners) across waves.

Self-esteem. Self-esteem was assessed with the 10-item RSE (Rosenberg, 1965).

Responses were measured on a 5-point scale (1 = *strongly disagree*; 5 = *strongly agree*), with $M = 4.12$ ($SD = 0.75$) averaged across waves. The alpha reliability ranged from .89 to .91 (male partners) across waves and was .92 (female partners) at each wave.

Attrition analyses. Data on study variables were available for 186 couples at Wave 1 and 173 couples at Wave 2. To examine the potential impact of attrition, we compared individuals who did versus did not participate in Wave 2 on study variables assessed at Wave 1. Male participants who dropped out reported lower life satisfaction ($M = 3.65$ vs. 5.08 ; $d = -1.18$) and higher depression ($M = 0.77$ vs. 0.54 ; $d = 0.59$); differences in self-esteem among male participants and differences in life satisfaction, depression, and self-esteem among female participants were nonsignificant. Given the relatively low dropout rate (i.e., 7%) and given that we used FIML for dealing with missing data, we concluded that attrition was not a critical issue in the MPI.

Procedure for the Statistical Analyses

The analyses were conducted using the Mplus 6.1 program (Muthén & Muthén, 2010). To deal with missing values, we used FIML estimation to fit models directly to the raw data (J. L. Schafer & Graham, 2002; Widaman, 2006). Model fit was assessed by the Tucker–Lewis index (TLI), the comparative fit index (CFI), and the root-mean-square error of approximation

(RMSEA), based on the recommendations of Hu and Bentler (1999) and MacCallum and Austin (2000). Hu and Bentler (1999) suggest that good fit is indicated by values greater than or equal to .95 for TLI and CFI, and less than or equal to .06 for RMSEA. To test for differences in model fit, we used the test of small difference in fit recommended by MacCallum, Browne, and Cai (2006, Program C).

For the analyses, we used a longitudinal version of the actor-partner interdependence model (Cook & Kenny, 2005; Kenny, Kashy, & Cook, 2006). The actor-partner interdependence model is particularly suited for the present analyses because it appropriately accounts for the nonindependence of data, when responses from both partners of couples are analyzed (Kenny & Cook, 1999; Townsend, 2012). Moreover, an important advantage is that the actor-partner interdependence model allows estimating within-person effects (called actor effects) and between-person effects (called partner effects) simultaneously, thereby mutually controlling the effects for each other. The key advantage of the longitudinal design of the model is that it allows controlling for each person's prior level of the outcome. Controlling for prior levels of the outcomes is important, because only then do partner effects inform about *change* in a construct due to the influence of the partner (e.g., Finkel, 1995; Little, Preacher, Selig, & Card, 2007).

In this research, all constructs were examined as latent variables, which produces more valid results by taking error variance into account (Cole & Preacher, 2014). We used item parcels as indicators because they produce more reliable latent variables than individual items (Little, Cunningham, Shahar, & Widaman, 2002). For each latent factor, we aggregated the items into three parcels. The error variances of each parcel were allowed to correlate across waves to control for bias due to parcel-specific variance (Cole & Maxwell, 2003). Metric invariance of the measures was established by constraining the loadings of the parcels to be equal across waves; as

indicated by the test of small difference in fit, these constraints did not significantly worsen model fit (see the supplemental materials, Table S6). In the analyses, we imposed metric invariance because this level of invariance is required for valid conclusions about the relations between factors such as correlations and regression coefficients (Schmitt & Kuljanin, 2008), corresponding to the effects of interests, that is, mutual influence and shared environment. For studies that covered more than one time interval (i.e., studies with more than two waves), the actor and partner effects were constrained to be equal across time intervals, which increases the validity of the estimates by aggregating the information across waves; again, these constraints did not significantly worsen model fit (see the supplemental materials, Table S7).⁴

For the meta-analytic computations, we used SPSS 20 and the SPSS macros written by Daniel B. Wilson (Lipsey & Wilson, 2001, Appendix D). All computations with effect sizes were made using Fisher's Z_r transformations. For computing the weighted mean effect sizes, we used fixed-effects models and study weights with $w = n - 3$ (Lipsey & Wilson, 2001; Raudenbush, 2009). Fixed-effect models rather than random-effects models were used because when the number of studies is small (i.e., five or less) random-effects models produce imprecise confidence intervals (Hedges & Vevea, 1998).

Results

Study-Level Analyses

Figure 1 provides a generic illustration of the longitudinal actor-partner interdependence model, for the example of a study with two waves of measurement. The model includes two predictor variables and two outcome variables (e.g., life satisfaction of the male and female partner at Time 1 and Time 2, respectively). The relations between the variables are modeled as

actor effects, partner effects, and within-wave correlations between the Time 1 variables and the unexplained variances of the Time 2 variables.

The model provides information on four important estimates. First, the correlation between the Time 1 variables captures the degree to which the partners are similar in the construct under investigation. Second, the actor effects indicate the stability of individual differences in the construct over time. Third, the partner effects indicate the influence of one partner on the other; if both men and women have significant partner effects on each other, then there is evidence of mutual influence between the partners. Fourth, given that the unexplained variances (i.e., residuals) at Time 2 account for the influence of factors that are not included in the model, the correlation between the unexplained variances captures the degree to which the partners are influenced in the same way by third variables, indicating the effect of shared environment (Caspi et al., 1992). As Caspi et al. (1992, p. 286) explain, “the correlation between the residuals should be zero, *unless there is some unspecified influence that is shared by members of the dyad*. If that is the case, the correlation between the residuals would be positive and significant.” In the context of the present research, the effects of mutual influence and shared environment are the key parameters of interest because these two processes account for codevelopment in the two partners.

We first tested whether actor and partner effects differed across gender by comparing the fit of a model in which the actor and partner effects are constrained to be equal across gender with the fit of a model in which the actor and partner effects are estimated freely for male and female partners. In all studies, cross-gender equality constraints did not significantly decrease model fit (Table 3). Consequently, we used these constraints in the remainder of the analyses. In all studies, the fit of the models was good (Table 4). Table 5 shows the estimates of similarity

effects (i.e., Time 1 correlation between partners), stability effects (i.e., actor effects), mutual influence effects (i.e., partner effects), and the effect of shared environment (i.e., the correlation between the unexplained variances at Time 2 and, if applicable, at later waves; for studies that included more than two waves, the correlations were averaged across waves). Across studies, most of the similarity effects, all stability effects, 8 out of 16 mutual influence effects, and 10 out of 16 shared environment effects were significant. All effects were in the expected direction (with the exception of the self-esteem partner effect for the IYFP, which was, however, nonsignificant).

Meta-Analytic Aggregation of the Effects

Next, we meta-analytically aggregated the effects across studies (Table 6). For all constructs, the similarity effect was significant, ranging from .13 (self-esteem) to .34 (negative affect). Not surprisingly, the stability effect was large and significant for all constructs, ranging from .62 to .81. The mutual influence effect was significant for life satisfaction, negative affect, positive affect, and depression, but not for self-esteem. The largest mutual influence effect emerged for life satisfaction (.12). The shared environment effect was significant for all constructs, ranging from .09 (positive affect) to .21 (negative affect).

Because of conceptual overlap between the measures of well-being used in the present research (i.e., life satisfaction, positive affect, negative affect, and depression), we also examined meta-analytic estimates when aggregating across well-being measures. Specifically, we first averaged the study-level estimates across measures of well-being and then meta-analytically aggregated the estimates across studies. Moreover, we tested whether the estimates differed significantly between well-being and self-esteem. When testing for the difference between estimates based on the same sample, it is incorrect to examine whether the confidence intervals

overlap; instead, it is the confidence interval of the *difference* that should be examined (Cumming, 2014). We therefore meta-analyzed the differences between well-being and self-esteem with regard to similarity, stability, mutual influence, and shared environment. The results of these analyses are shown at the bottom of Table 6. For well-being, the findings suggest that all effects were significant. Importantly, there was evidence of mutual influence between partners (.06) and shared environmental effects (.13). Moreover, the estimates for well-being differed significantly from the estimates for self-esteem, except with regard to shared environment. The findings suggest that partners are more similar in their well-being than in their self-esteem (specifically, the effects were .26 vs. .13). Moreover, the findings indicate that mutual influence is specific to well-being (i.e., the difference was significant, in addition to the fact that the effect was nonsignificant for self-esteem). Finally, the stability of well-being was significantly lower than that of self-esteem.

Discussion

In this research, we tested two explanations for why romantic partners develop in similar ways in their well-being and self-esteem. Specifically, to what degree is codevelopment due to mutual influence between partners or to the effects of shared environment? We used longitudinal data from five samples, which included dating, cohabiting, and married couples and which covered a broad age range, from young adulthood to old age. For the analyses, we used the longitudinal actor-partner interdependence model. After conducting the study-level analyses, we meta-analytically aggregated the findings across studies. The results showed significant mutual influence between partners in life satisfaction, positive affect, negative affect, and depression, and significant effects of shared environment on life satisfaction, positive affect, negative affect, depression, and self-esteem. The findings suggest that both mutual influence and shared

environment account for codevelopment of well-being in romantic partners. In contrast, only shared environment but not mutual influence contributes to the partners' codevelopment of self-esteem.

Implications of the Findings

With regard to well-being, the present research suggests that one reason for codevelopment is that relationship partners mutually influence each other. Thus, the present research supports the perspectives of Hatfield et al.'s (1993) emotional contagion theory, Coyne's (1976) interactional theory of depression, and Kelley and Thibaut's (1978) interdependence theory. It is important to note that the studies included in the present research covered periods of several years (with time lags between waves ranging from 6 months to 3 years), providing evidence that mutual influence between partners clearly goes beyond daily influences and accounts for long-term effects on well-being. Although all of the theories cited above predict that the well-being of one partner contributes to the well-being of the other partner and vice versa, interdependence theory might provide the best basis to extend the understanding of long-term mutual influence in the well-being of romantic partners. In particular, interdependence theory states that close relationships differ with regard to the degree of dependence as well as the mutuality of dependence (Rusbult & Buunk, 1993), suggesting possible moderators of the strength of partner effects in well-being. Rusbult's investment model of commitment (Rusbult & Buunk, 1993) builds on interdependence theory and proposes that people's feelings of commitment to a relationship capture their degree of dependence. This perspective suggests that when a person's commitment to the relationship is strong, his or her level of well-being should be strongly influenced by the partner's well-being; similarly, when a person's commitment is weak, partner effects on this person should be small or nonexistent. In

future research, it would be highly interesting to explore these hypotheses in more detail. In addition to mutual influence, the present research suggests that shared environment is another reason for why the well-being of partners develops in similar ways. For each measure of well-being, as well as when aggregating across measures of well-being, the effects of shared environment were significant (over and above the effects of mutual influence) and explained why the well-being of romantic partners changes in similar directions.

With regard to self-esteem, the present findings did not show evidence of mutual influence between partners. Across five datasets with an overall sample size of 2,058 couples, the effect size was virtually zero (.01). A possible explanation for this finding is that self-esteem is a more trait-like personality characteristic compared to well-being. Whereas research suggests that the rank-order stability of self-esteem is comparable to key personality characteristics such as the Big Five (Kuster & Orth, 2013; Roberts & DelVecchio, 2000; Trzesniewski, Donnellan, & Robins, 2003; Wagner, Lüdtkke, & Trautwein, 2016), the rank-order stability is lower for measures of well-being (Anusic & Schimmack, 2016; Lucas & Donnellan, 2007; Watson, 2004; Watson & Humrichouse, 2006). Correspondingly, in the present research, the stability effect of self-esteem was large (.81), and significantly larger than the stability effect of well-being (.68, when averaging across measures of well-being). Thus, the present findings do not suggest that there is mutual influence in self-esteem. However, self-esteem might indirectly influence the well-being of relationship partners. A growing body of research indicates that self-esteem has prospective effects on the person's level of well-being (Orth & Robins, 2013; Orth et al., 2012; Sowislo & Orth, 2013; Steiger, Fend, & Allemand, 2015), which in turn might affect the partner's well-being. Moreover, the meta-analytic aggregation suggested that shared environment had a significant effect on the self-esteem of romantic partners, contributing to

codevelopment of self-esteem in the two partners. Interestingly, the effect of shared environment on self-esteem (.11) was not significantly smaller than the effect on well-being (.13, when averaging across measures of well-being). Thus, despite evidence that self-esteem is more trait-like than well-being, the findings suggest that the self-esteem of romantic partners is influenced by shared environmental factors as strongly as well-being.

In all five studies and for all constructs examined, gender differences in the influences between partners were nonsignificant, despite relatively large sample sizes. In principle, it might be possible to speculate why one of the genders might be more strongly influenced by a relationship partner than the other gender. For example, men and women might differ in the normative level of socially dominant behavior or they might differ in the degree to which social relationships influence the person (compare, e.g., the concept of contingencies of self-esteem by Crocker & Wolfe, 2001). However, the present results clearly suggest that the size of partner effects does not depend on the direction of the effects (i.e., from the female partner to the male partner or vice versa). Put differently, mutual influence between male and female partners was truly reciprocal.

According to Cohen (1992), the effects of mutual influence in measures of well-being were small (ranging from .05 to .12) and the effects of shared environment were of small to medium size (ranging from .09 to .21). However, we argue that, for several reasons, the effects might be more important than suggested by Cohen's guidelines for small, medium, and large correlations. First, whereas Cohen's guidelines are typically applied to cross-sectional correlations, in the present research the effects are based on longitudinal data, controlling for prior levels of the outcomes (i.e., well-being and self-esteem at the preceding assessment). Second, in addition to controlling for prior levels, the effects of mutual influence and shared

environment are mutually controlled for each other, providing information about the size of unique effects of these influences on codevelopment. Third, the effects of mutual influence and shared environment cannot be explained by shared method variance (Podsakoff, MacKenzie, & Podsakoff, 2012), because the effects are based on data from two different sources (i.e., the two partners). Thus, whereas in many research situations effect sizes can be inflated by shared method variance, in the present research this possibility is ruled out by the design of the analyses.

Besides the effects of mutual influence and shared environment, the present research provides evidence on the degree to which relationship partners are similar in their well-being and self-esteem. When meta-analytically aggregated across studies, similarity was largest for negative affect (.34) and life satisfaction (.30), corresponding to medium-sized effects according to Cohen (1992). The smallest similarity effect emerged for self-esteem (.13). Overall, these similarity effects correspond to findings from previous studies (e.g., Bookwala & Schulz, 1996; Du Rocher Schudlich et al., 2011; Erol & Orth, 2013; Watson et al., 2004). An important question is whether similarity is beneficial for relationships. Theory suggests that similarity contributes to relationship quality, because it increases relationship functioning and reduces conflicts in daily life, and because similar partners share more similar emotional responses (Dyrenforth, Kashy, Donnellan, & Lucas, 2010). Nevertheless, the available evidence on the effects on well-being and self-esteem is mixed, suggesting that similarity has, at most, small effects on the partners' satisfaction with the relationship (Anderson et al., 2003; Donnellan, Assad, Robins, & Conger, 2007; Erol & Orth, 2014; Robins, Caspi, & Moffitt, 2000; for a review, see Erol & Orth, 2016).

Limitations and Future Directions

A limitation of the present research is that all samples were from Western countries. Therefore, future research should replicate the findings in samples from other cultural contexts, such as African, Asian, and Latin American countries (Arnett, 2008; Henrich, Heine, & Norenzayan, 2010). For example, given that interaction and communication styles among partners vary as a function of cultural context (e.g., Williamson et al., 2012), it is possible that cultures differ in the typical degree of codevelopment among relationship partners and in the processes that account for codevelopment.

Another limitation is that the present research did not examine the mechanisms that explain mutual influence between relationship partners. Importantly, it is possible that different mechanisms account for the effects on different constructs. For example, research suggests that higher life satisfaction leads to more functional providing of social support (Coffman & Gilligan, 2002), which might influence the partner's life satisfaction. Also, research indicates that depressed individuals tend to excessively ask their partner for reassurance of their worth, which might negatively affect their partner's well-being (Katz et al., 1999). Future research should test hypotheses on mediating mechanisms that are tailored specifically to the construct under investigation.

In the present research, the effects of shared environment were not measured directly, but were assessed indirectly on the basis of residual correlations. Therefore, it may be useful to discuss what the correlation between the residuals captures and what does it not capture. First, even if there is unexplained variance (i.e., residual variance) in the constructs at Time 2 and later assessments, these residuals are not necessarily correlated between partners. Thus, it is an empirical question to what degree the residuals are similar across partners. Although our findings do suggest that the residuals are correlated between partners, these correlations could have been

zero. Second, at each assessment, similarity between the partners can be explained by an important source of similarity that is distinct from mutual influence and shared environment, specifically, by similarity at Time 1. Thus, any observed similarity at Time 2 (and later assessments) is not only explained by the cross-lagged partner effects (i.e., mutual influence) and by the correlation between the residuals (i.e., shared environment), but also by the Time 1 correlation (i.e., similarity at the beginning of the study). The reason is that the Time 1 similarity is carried forward by the stability of the constructs, following the tracing rules of structural equation modeling (e.g., Kenny, 1979). In other words, the correlation between the residuals is not confounded by the degree of similarity that existed already at Time 1 (i.e., this source of similarity is controlled for).

However, an important question is whether the correlation between the residuals can indeed be interpreted as an indicator of the effects of shared environment. In particular, given that relationship partners are frequently of about the same age, the residual correlations could be artificially inflated simply because of similar age-related changes. In fact, as mentioned above, well-being and self-esteem show normative changes during adulthood (e.g., Carstensen et al., 2011; Mroczek & Spiro, 2005; Orth & Robins, 2014). However, we argue that age is not an explanation that invalidates the interpretation that residual correlations are due to shared environmental effects. Although age may provide a statistical explanation of developmental change, age is an “empty variable” (Neugarten, 1977, p. 633) that is only a proxy for the underlying biological, social, and psychological factors influencing development (Dixon, 2011). Theory suggests that age-related changes in well-being and self-esteem during adulthood are influenced by factors such as adopting new social roles (e.g., roles of employee, supervisor, spouse, parent, and so on), life events and life transitions, as well as changes in people’s goals

(Hutteman, Hennecke, Orth, Reitz, & Specht, 2014; Roberts & Wood, 2006; Roberts, Wood, & Caspi, 2008). Moreover, although there are age-related patterns of development, most individuals do not follow the normative pattern of development exactly, as indicated by the large variability of individual trajectories with regard to levels, slopes, and turning points (e.g., Charles et al., 2001; Orth et al., 2010). Consequently, if change is correlated within couples (controlling for mutual influence between partners), this suggests that influences shared by both partners account for the association (Caspi et al., 1992).

Nevertheless, a qualification is that environmental influences that are shared by almost every member of a specific age group could affect the residual correlation when the sample is heterogeneous with regard to age, as is true for all samples examined in this research. Thus, the residual correlation could be influenced not only by environmental factors that are specific to a couple (such as life events, moving to another city, and changes in the social network, to name but a few) but also by environmental factors that are shared by members of a specific age group (such as having young children, the empty nest situation, and retirement). In the present research, the sample sizes were not sufficiently large for testing the models in age-homogeneous subsamples (e.g., in 5-year age groups). Thus, future research should test whether the present findings hold when using samples that are homogeneous with regard to age. Another qualification is that partners may show some degree of genetic similarity resulting from assortative mating, as reflected for example in the finding that partners tend to show some similarity in temperament traits (Watson et al., 2004). However, we argue that genetic factors contribute at most in minor ways to residual correlations at Time 2 and later assessments. The reason is that, even if partners show some degree of genetic similarity, most of this similarity is captured already by the Time 1 correlation between partners and, as described above, the effect

of the Time 1 correlation is controlled for in the residual correlation. Strictly speaking, however, the model does not rule out the possibility that genetic similarity between partners could lead to similar changes over time in relationship partners, thereby increasing residual correlations. In sum, it is possible that the residual correlation overestimates the size of shared environmental effects, but we believe that such a bias is small. Future research should seek to directly test for the effects of shared environment and to identify the key factors in the shared environment of relationship partners.

The between-study differences in coefficients raise the important question of whether sample characteristics such as age, length of relationship, country, or ethnicity account for systematic differences in the effect sizes. For example, a hypothesis could be that mutual influence between partners is larger at the early stages of romantic relationships (compared to long-standing relationships), because the influence of the partner is still relatively new and, consequently, might be more powerful. Moreover, as mentioned earlier, it is possible that the typical degree of mutual influence differs across countries and cultural contexts, for example because social and religious norms influence the typical degree of closeness in romantic relationships, which might moderate the degree of mutual influence between partners (Rusbult & Buunk, 1993). However, in the present research, the number of studies (ranging from 2 to 5, depending on the analysis) was too small to provide for sufficient power and reliability in meta-analytic tests of moderators of the effect sizes (e.g., Cooper, Hedges, & Valentine, 2009; Goh et al., 2016). Thus, in future research it would be worthwhile to test for moderators of mutual influence and shared environmental effects once a larger number of dyadic studies is available for meta-analysis. Nevertheless, we note that with regard to mutual influence and shared environment (i.e., the central coefficients of interest in this research), nearly all homogeneity

statistics were nonsignificant, suggesting that the observed between-study differences cannot be attributed to between-study sampling error but only to within-study sampling error (Cooper et al., 2009; Lipsey & Wilson, 2001). Thus, with regard to the five studies examined in the present research, the results did not suggest that the coefficients differed more strongly across studies than would be expected by chance. Given these findings, we believe that any observed differences between studies should not be interpreted (e.g., why there was significant mutual influence in self-esteem in the MPI but not in the other studies, a finding that could be simply due to chance), but that the conclusions should be based only on the findings combined across studies (i.e., the meta-analytical estimates).

A strength of the present research is that we used data from five independent longitudinal studies. The studies were heterogeneous with regard to methodological characteristics (such as sampling procedure, assessment methods, and time interval between assessments), which strengthens the generalizability of the combined findings. Also, the overall sample size across studies was large (i.e., 2,058 couples), increasing the precision of the estimates. Moreover, most of the measures used were established multi-item measures of the constructs, and all constructs were examined as latent variables, which strengthens the validity of the findings. Thus, the present research provides robust evidence on the effects of mutual influence and shared environment on codevelopment of well-being and self-esteem in romantic partners.

Conclusion

The findings suggest that both mutual influence and shared environment contribute to the codevelopment of well-being among relationship partners, whereas, for self-esteem, only shared environment but not mutual influence explains why partners develop in similar ways. The findings have important practical implications because they show that the level of well-being of a

person's relationship partner has long-term influence on the person's own well-being.

Consequently, interventions aimed at improving the well-being of an individual might be more effective if the partner's well-being is taken into account.

References

- Allison, P. D. (2003). Missing data techniques for structural equation modeling. *Journal of Abnormal Psychology, 112*, 545-557. doi: 10.1037/0021-843X.112.4.545
- Anderson, C., Keltner, D., & John, O. P. (2003). Emotional convergence between people over time. *Journal of Personality and Social Psychology, 84*, 1054-1068. doi: 10.1037/0022-3514.84.5.1054
- Anusic, I., & Schimmack, U. (2016). Stability and change of personality traits, self-esteem, and well-being: Introducing the meta-analytic stability and change model of retest correlations. *Journal of Personality and Social Psychology, 110*, 766-781. doi: 10.1037/pspp0000066
- Arnett, J. J. (2008). The neglected 95%: Why American psychology needs to become less American. *American Psychologist, 63*, 602-614. doi: 10.1037/0003-066X.63.7.602
- Barnett, R. C. (1993). *Longitudinal Study of Dual-Earner Couples, 1989-1992 [Data file and codebook]*. Cambridge, MA: Murray Research Archive [Distributor].
- Bengtson, V. L. (2009). *Longitudinal Study of Generations, 1971, 1985, 1988, 1991, 1994, 1997, 2000 [Data file and codebook]*. Ann Arbor, MI: Inter-University Consortium for Political and Social Research [Distributor]. doi: 10.3886/ICPSR22100.v2
- Berman, A. L. (2009). Depression and suicide. In I. H. Gotlib & C. L. Hammen (Eds.), *Handbook of depression* (pp. 510-530). New York, NY: Guilford.
- Bookwala, J., & Schulz, R. (1996). Spousal similarity in subjective well-being: The Cardiovascular Health Study. *Psychology and Aging, 11*, 582-590. doi: 10.1037/0882-7974.11.4.582
- Bradburn, N. M. (1969). *The structure of psychological well-being*. Chicago, IL: Aldine.

- Busseri, M. A., & Sadava, S. W. (2011). A review of the tripartite structure of subjective well-being: Implications for conceptualization, operationalization, analysis, and synthesis. *Personality and Social Psychology Review, 15*, 290-314. doi: 10.1177/1088868310391271
- Carstensen, L. L., Turan, B., Scheibe, S., Ram, N., Ersner-Hershfield, H., Samanez-Larkin, G. R., . . . Nesselroade, J. R. (2011). Emotional experience improves with age: Evidence based on over 10 years of experience sampling. *Psychology and Aging, 26*, 21-33. doi: 10.1037/a0021285
- Caspi, A., Herbener, E. S., & Ozer, D. J. (1992). Shared experiences and the similarity of personalities: A longitudinal study of married couples. *Journal of Personality and Social Psychology, 62*, 281-291. doi: 10.1037/0022-3514.62.2.281
- Charles, S. T., Reynolds, C. A., & Gatz, M. (2001). Age-related differences and change in positive and negative affect over 23 years. *Journal of Personality and Social Psychology, 80*, 136-151. doi: 10.1037/0022-3514.80.1.136
- Coffman, D. L., & Gilligan, T. D. (2002). Social support, stress, and self-efficacy: Effects on students' satisfaction. *College Student Retention, 4*, 53-66. doi: 10.2190/BV7X-F87X-2MXL-2B3L
- Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*, 155-159. doi: 10.1037/0033-2909.112.1.155
- Cole, D. A., & Maxwell, S. E. (2003). Testing mediational models with longitudinal data: Questions and tips in the use of structural equation modeling. *Journal of Abnormal Psychology, 112*, 558-577. doi: 10.1037/0021-843X.112.4.558

- Cole, D. A., & Preacher, K. J. (2014). Manifest variable path analysis: Potentially serious and misleading consequences due to uncorrected measurement error. *Psychological Methods, 19*, 300-315. doi: 10.1037/a0033805
- Conger, R. D., Lasley, P., Lorenz, F. O., Simons, R., Whitbeck, L. B., Elder, G. H., & Norem, R. (2011). *Iowa Youth and Families Project, 1989-1992 [Data file and codebook]*. Ann Arbor, MI: Inter-University Consortium for Political and Social Research [Distributor]. doi: 10.3886/ICPSR26721.v2
- Cook, W. L., & Kenny, D. A. (2005). The actor-partner interdependence model: A model of bidirectional effects in developmental studies. *International Journal of Behavioral Development, 29*, 101-109. doi: 10.1080/01650250444000405
- Cooper, H., Hedges, L. V., & Valentine, J. C. (Eds.). (2009). *The handbook of research synthesis and meta-analysis*. New York, NY: Russell Sage Foundation.
- Coyne, J. C. (1976). Depression and the response of others. *Journal of Abnormal Psychology, 85*, 186-193. doi: 10.1037/0021-843X.85.2.186
- Crocker, J., & Wolfe, C. T. (2001). Contingencies of self-worth. *Psychological Review, 108*, 593-623. doi: 10.1037/0033-295X.108.3.593
- Cumming, G. (2014). The new statistics: Why and how. *Psychological Science, 25*, 7-29. doi: 10.1177/0956797613504966
- Demerouti, E., Bakker, A. B., & Schaufeli, W. B. (2005). Spillover and crossover of exhaustion and life satisfaction among dual-earner parents. *Journal of Vocational Behavior, 67*, 266-289. doi: 10.1016/j.jvb.2004.07.001

- Derogatis, L. R., & Savitz, K. L. (1999). The SCL-90-R, Brief Symptom Inventory, and matching clinical rating scales. In M. E. Maruish (Ed.), *The use of psychological testing for treatment planning and outcomes assessment* (pp. 679-724). Mahwah, NJ: Erlbaum.
- Diener, E. (1984). Subjective well-being. *Psychological Bulletin*, *95*, 542-575. doi: 10.1037/0033-2909.95.3.542
- Diener, E. (2000). Subjective well-being: The science of happiness and a proposal for a national index. *American Psychologist*, *55*, 34-43. doi: 10.1037/0003-066X.55.1.34
- Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The Satisfaction With Life Scale. *Journal of Personality Assessment*, *49*, 71-75. doi: 10.1207/s15327752jpa4901_13
- Diener, E., Suh, E. M., Lucas, R. E., & Smith, H. L. (1999). Subjective well-being: Three decades of progress. *Psychological Bulletin*, *125*, 276-302. doi: 10.1037/0033-2909.125.2.276
- Diener, M. L., & Diener McGavran, M. B. (2008). What makes people happy? A developmental approach to the literature on family relationships and well-being. In M. Eid & R. J. Larsen (Eds.), *The science of subjective well-being* (pp. 347-375). New York, NY: Guilford.
- Dixon, R. A. (2011). Enduring theoretical themes in psychological aging: Derivation, functions, perspectives, and opportunities. In K. W. Schaie & S. L. Willis (Eds.), *Handbook of the psychology of aging* (pp. 3-23). Amsterdam: Elsevier.
- Donnellan, M. B., Assad, K. K., Robins, R. W., & Conger, R. D. (2007). Do negative interactions mediate the effects of negative emotionality, communal positive emotionality, and constraint on relationship satisfaction? *Journal of Social and Personal Relationships*, *24*, 557-573. doi: 10.1177/0265407507079249

- Du Rocher Schudlich, T. D., Papp, L. M., & Cummings, E. M. (2011). Relations between spouses' depressive symptoms and marital conflict: A longitudinal investigation of the role of conflict resolution styles. *Journal of Family Psychology, 25*, 531-540. doi: 10.1037/a0024216
- Dyrenforth, P. S., Kashy, D. A., Donnellan, M. B., & Lucas, R. E. (2010). Predicting relationship and life satisfaction from personality in nationally representative samples from three countries: The relative importance of actor, partner, and similarity effects. *Journal of Personality and Social Psychology, 99*, 690-702. doi: 10.1037/a0020385
- Eaton, W. W., Smith, C., Ybarra, M., Muntaner, C., & Tien, A. (2004). Center for Epidemiologic Studies Depression Scale: Review and revision (CESD and CESD-R). In M. E. Maruish (Ed.), *The use of psychological testing for treatment planning and outcomes assessment: Volume 3 (Instruments for adults)* (pp. 363-377). Mahwah, NJ: Erlbaum.
- Erol, R. Y., & Orth, U. (2013). Actor and partner effects of self-esteem on relationship satisfaction and the mediating role of secure attachment between the partners. *Journal of Research in Personality, 47*, 26-35. doi: 10.1016/j.jrp.2012.11.003
- Erol, R. Y., & Orth, U. (2014). Development of self-esteem and relationship satisfaction in couples: Two longitudinal studies. *Developmental Psychology, 50*, 2291-2303. doi: 10.1037/a0037370
- Erol, R. Y., & Orth, U. (2016). Self-esteem and the quality of romantic relationships. *European Psychologist, 21*, 274-283. doi: 10.1027/1016-9040/a000259
- Finkel, S. E. (1995). *Causal analysis with panel data*. Thousand Oaks, CA: Sage.
- Furler, K., Gomez, V., & Grob, A. (2014). Personality perceptions and relationship satisfaction in couples. *Journal of Research in Personality, 50*, 33-41. doi: 10.1016/j.jrp.2014.02.003

Gaunt, R. (2006). Couple similarity and marital satisfaction: Are similar spouses happier?

Journal of Personality, 74, 1402-1420. doi: 10.1111/j.1467-6494.2006.00414.x

Gerstorf, D., Hoppmann, C. A., Kadlec, K. M., & McArdle, J. J. (2009). Memory and depressive symptoms are dynamically linked among married couples: Longitudinal evidence from the AHEAD study. *Developmental Psychology, 45*, 1595-1610. doi: 10.1037/a0016346

Goh, J. X., Hall, J. A., & Rosenthal, R. (2016). Mini meta-analysis of your own studies: Some arguments on why and a primer on how. *Social and Personality Psychology Compass, 10*, 535-549. doi: 10.1111/spc3.12267

Gonzaga, G. C., Campos, B., & Bradbury, T. (2007). Similarity, convergence, and relationship satisfaction in dating and married couples. *Journal of Personality and Social Psychology, 93*, 34-48. doi: 10.1037/0022-3514.93.1.34

Harding, S. D. (1982). Psychological well-being in Great Britain: An evaluation of the Bradburn Affect Balance Scale. *Personality and Individual Differences, 3*, 167-175. doi: 10.1016/0191-8869(82)90031-9

Hatfield, E., Cacioppo, J. T., & Rapson, R. L. (1993). Emotional contagion. *Current Directions in Psychological Science, 2*, 96-99. doi: 10.1111/1467-8721.ep10770953

Hedges, L. V., & Vevea, J. L. (1998). Fixed- and random-effects models in meta-analysis.

Psychological Methods, 3, 486-504. doi: 10.1037/1082-989X.3.4.486

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world?

Behavioral and Brain Sciences, 33, 61-83. doi: 10.1017/S0140525X0999152X

Hoppmann, C. A., Gerstorf, D., Willis, S. L., & Schaie, K. W. (2011). Spousal interrelations in happiness in the Seattle Longitudinal Study: Considerable similarities in levels and change over time. *Developmental Psychology, 47*, 1-8. doi: 10.1037/a0020788

- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1-55. doi: 10.1080/10705519909540118
- Hutteman, R., Hennecke, M., Orth, U., Reitz, A. K., & Specht, J. (2014). Developmental tasks as a framework to study personality development in adulthood and old age. *European Journal of Personality, 28*, 267-278. doi: 10.1002/per.1959
- Joiner, T. E., & Katz, J. (1999). Contagion of depressive symptoms and mood: Meta-analytic review and explanations from cognitive, behavioral, and interpersonal viewpoints. *Clinical Psychology: Science and Practice, 6*, 149-164. doi: 10.1093/clipsy.6.2.149
- Katz, J., Beach, S. R. H., & Joiner, T. E. (1999). Contagious depression in dating couples. *Journal of Social and Clinical Psychology, 18*, 1-13. doi: 10.1521/jscp.1999.18.1.1
- Kelley, H. H., & Thibaut, J. W. (1978). *Interpersonal relations: A theory of interdependence*. New York, NY: Wiley.
- Kenny, D. A. (1979). *Correlation and causality*. New York, NY: Wiley.
- Kenny, D. A., & Cook, W. (1999). Partner effects in relationship research: Conceptual issues, analytic difficulties, and illustrations. *Personal Relationships, 6*, 433-448. doi: 10.1111/j.1475-6811.1999.tb00202.x
- Kenny, D. A., Kashy, D. A., & Cook, W. L. (2006). *Dyadic data analysis*. New York, NY: Guilford.
- Kessler, R. C., Berglund, P., Demler, O., Jin, R., Merikangas, K. R., & Walters, E. E. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry, 62*, 593-602. doi: 10.1001/archpsyc.62.6.593

- Kim, K. A., & Mueller, D. J. (2001). To balance or not to balance: Confirmatory factor analysis of the Affect-Balance Scale. *Journal of Happiness Studies*, 2, 289-306. doi: 10.1023/A:1013519931082
- Kouros, C. D., & Cummings, E. M. (2010). Longitudinal associations between husbands' and wives' depressive symptoms. *Journal of Marriage and Family*, 72, 135-147. doi: 10.1111/j.1741-3737.2009.00688.x
- Kuster, F., & Orth, U. (2013). The long-term stability of self-esteem: Its time-dependent decay and nonzero asymptote. *Personality and Social Psychology Bulletin*, 39, 677-690. doi: 10.1177/0146167213480189
- Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.
- Little, T. D., Cunningham, W. A., Shahar, G., & Widaman, K. F. (2002). To parcel or not to parcel: Exploring the question, weighing the merits. *Structural Equation Modeling*, 9, 151-173. doi: 10.1207/S15328007SEM0902_1
- Little, T. D., Preacher, K. J., Selig, J. P., & Card, N. A. (2007). New developments in latent variable panel analyses of longitudinal data. *International Journal of Behavioral Development*, 31, 357-365. doi: 10.1177/0165025407077757
- Lucas, R. E., Diener, E., & Suh, E. (1996). Discriminant validity of well-being measures. *Journal of Personality and Social Psychology*, 71, 616-628. doi: 10.1037/0022-3514.71.3.616
- Lucas, R. E., & Donnellan, M. B. (2007). How stable is happiness? Using the STARTS model to estimate the stability of life satisfaction. *Journal of Research in Personality*, 41, 1091-1098. doi: 10.1016/j.jrp.2006.11.005

- Luciano, E. C., & Orth, U. (2017). Transitions in romantic relationships and development of self-esteem. *Journal of Personality and Social Psychology, 112*, 307-328. doi: 10.1037/pspp0000109
- Luhmann, M., Weiss, P., Hosoya, G., & Eid, M. (2014). Honey, I got fired! A longitudinal dyadic analysis of the effect of unemployment on life satisfaction in couples. *Journal of Personality and Social Psychology, 107*, 163-180. doi: 10.1037/a0036394
- Luo, S., & Klohnen, E. C. (2005). Assortative mating and marital quality in newlyweds: A couple-centered approach. *Journal of Personality and Social Psychology, 88*, 304-326. doi: 10.1037/0022-3514.88.2.304
- MacCallum, R. C., & Austin, J. T. (2000). Applications of structural equation modeling in psychological research. *Annual Review of Psychology, 51*, 201-226. doi: 10.1146/annurev.psych.51.1.201
- MacCallum, R. C., Browne, M. W., & Cai, L. (2006). Testing differences between nested covariance structure models: Power analysis and null hypotheses. *Psychological Methods, 11*, 19-35. doi: 10.1037/1082-989X.11.1.19
- Meyler, D., Stimpson, J. P., & Peek, M. K. (2007). Health concordance within couples: A systematic review. *Social Science and Medicine, 64*, 2297-2310. doi: 10.1016/j.socscimed.2007.02.007
- Mroczek, D. K., & Kolarz, C. M. (1998). The effect of age on positive and negative affect: A developmental perspective on happiness. *Journal of Personality and Social Psychology, 75*, 1333-1349. doi: 10.1037/0022-3514.75.5.1333

- Mroczek, D. K., & Spiro, A. (2005). Change in life satisfaction during adulthood: Findings from the Veterans Affairs Normative Aging Study. *Journal of Personality and Social Psychology, 88*, 189-202. doi: 10.1037/0022-3514.88.1.189
- Murdock, K. W., Lovejoy, M. C., & Oddi, K. B. (2014). An actor-partner interdependence analysis of associations between affect and parenting behavior among couples. *Family Process, 53*, 120-130. doi: 10.1111/famp.12059
- Murray, S. L., Holmes, J. G., & Griffin, D. W. (2000). Self-esteem and the quest for felt security: How perceived regard regulates attachment processes. *Journal of Personality and Social Psychology, 78*, 478-498. doi: 10.1037/0022-3514.78.3.478
- Muthén, L. K., & Muthén, B. O. (2010). *Mplus user's guide: Sixth edition*. Los Angeles, CA: Muthén and Muthén.
- Neugarten, B. L. (1977). Personality and aging. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (pp. 626-649). New York, NY: Van Nostrand Reinhold.
- Orth, U. (2013). How large are actor and partner effects of personality on relationship satisfaction? The importance of controlling for shared method variance. *Personality and Social Psychology Bulletin, 39*, 1359-1372. doi: 10.1177/0146167213492429
- Orth, U., Maes, J., & Schmitt, M. (2015). Self-esteem development across the life span: A longitudinal study with a large sample from Germany. *Developmental Psychology, 51*, 248-259. doi: 10.1037/a0038481
- Orth, U., & Robins, R. W. (2013). Understanding the link between low self-esteem and depression. *Current Directions in Psychological Science, 22*, 455-460. doi: 10.1177/0963721413492763

- Orth, U., & Robins, R. W. (2014). The development of self-esteem. *Current Directions in Psychological Science*, 23, 381-387. doi: 10.1177/0963721414547414
- Orth, U., Robins, R. W., & Widaman, K. F. (2012). Life-span development of self-esteem and its effects on important life outcomes. *Journal of Personality and Social Psychology*, 102, 1271-1288. doi: 10.1037/a0025558
- Orth, U., Trzesniewski, K. H., & Robins, R. W. (2010). Self-esteem development from young adulthood to old age: A cohort-sequential longitudinal study. *Journal of Personality and Social Psychology*, 98, 645-658. doi: 10.1037/a0018769
- Peek, M. K., Stimpson, J. P., Townsend, A. L., & Markides, K. S. (2006). Well-being in older Mexican American spouses. *The Gerontologist*, 46, 258-265. doi: 10.1093/geront/46.2.258
- Podsakoff, P. M., MacKenzie, S. B., & Podsakoff, N. P. (2012). Sources of method bias in social science research and recommendations on how to control it. *Annual Review of Psychology*, 63, 539-569. doi: 10.1146/annurev-psych-120710-100452
- Radloff, L. S. (1977). The CES-D Scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement*, 1, 385-401. doi: 10.1177/014662167700100306
- Raudenbush, S. W. (2009). Analyzing effect sizes: Random-effects models. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (pp. 295-315). New York, NY: Russell Sage Foundation.
- Raykov, T., Dimitrov, D. M., & Asparouhov, T. (2010). Evaluation of scale reliability with binary measures using latent variable modeling. *Structural Equation Modeling*, 17, 265-279. doi: 10.1080/10705511003659417

- Roberts, B. W., & DelVecchio, W. F. (2000). The rank-order consistency of personality traits from childhood to old age: A quantitative review of longitudinal studies. *Psychological Bulletin, 126*, 3-25. doi: 10.1037/0033-2909.126.1.3
- Roberts, B. W., & Wood, D. (2006). Personality development in the context of the neo-socioanalytic model of personality. In D. K. Mroczek & T. D. Little (Eds.), *Handbook of personality development* (pp. 11-39). Mahwah, NJ: Erlbaum.
- Roberts, B. W., Wood, D., & Caspi, A. (2008). The development of personality traits in adulthood. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (pp. 375-398). New York, NY: Guilford.
- Robins, R. W., Caspi, A., & Moffitt, T. E. (2000). Two personalities, one relationship: Both partners' personality traits shape the quality of their relationship. *Journal of Personality and Social Psychology, 79*, 251-259. doi: 10.1037/0022-3514.79.2.251
- Robins, R. W., Hendin, H. M., & Trzesniewski, K. H. (2001). Measuring global self-esteem: Construct validation of a single-item measure and the Rosenberg Self-Esteem Scale. *Personality and Social Psychology Bulletin, 27*, 151-161. doi: 10.1177/0146167201272002
- Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.
- Rusbult, C. E., & Buunk, B. P. (1993). Commitment processes in close relationships: An interdependence analysis. *Journal of Social and Personal Relationships, 10*, 175-204. doi: 10.1177/026540759301000202

- Rusbult, C. E., & Van Lange, P. A. M. (2003). Interdependence, interaction, and relationships. *Annual Review of Psychology, 54*, 351-375. doi: 10.1146/annurev.psych.54.101601.145059
- Saxbe, D., & Repetti, R. L. (2010). For better or worse? Coregulation of couples' cortisol levels and mood states. *Journal of Personality and Social Psychology, 98*, 92-103. doi: 10.1037/a0016959
- Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods, 7*, 147-177. doi: 10.1037/1082-989X.7.2.147
- Schafer, R. B., & Keith, P. M. (1992). Self-esteem agreement in the marital relationship. *Journal of Social Psychology, 132*, 5-9. doi: 10.1080/00224545.1992.9924683
- 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. doi: 10.1016/j.jrp.2014.08.007
- Schimmack, U., & Lucas, R. E. (2010). Environmental influences on well-being: A dyadic latent panel analysis of spousal similarity. *Social Indicators Research, 98*, 1-21. doi: 10.1007/s11205-009-9516-8
- Schmitt, N., & Kuljanin, G. (2008). Measurement invariance: Review of practice and implications. *Human Resource Management Review, 18*, 210-222. doi: 10.1016/j.hrmr.2008.03.003
- Schumm, W. R., Figley, C. R., & Fuhs, N. N. (1980). Similarity in self-esteem as a function of duration of marriage among student couples. *Psychological Reports, 47*, 365-366. doi: 10.2466/pr0.1980.47.2.365

- Segrin, C., Badger, T., Dorros, S. M., Meek, P., & Lopez, A. M. (2007). Interdependent anxiety and psychological distress in women with breast cancer and their partners. *Psycho-Oncology, 16*, 634-643. doi: 10.1002/pon.1111
- Segrin, C., Badger, T. A., & Harrington, J. (2012). Interdependent psychological quality of life in dyads adjusting to prostate cancer. *Health Psychology, 31*, 70-79. doi: 10.1037/a0025394
- Segrin, C., Badger, T. A., Meek, P., Lopez, A. M., Bonham, E., & Sieger, A. (2005). Dyadic interdependence on affect and quality-of-life trajectories among women with breast cancer and their partners. *Journal of Social and Personal Relationships, 22*, 673-689. doi: 10.1177/0265407505056443
- Segrin, C., Powell, H. L., Givertz, M., & Brackin, A. (2003). Symptoms of depression, relational quality, and loneliness in dating relationships. *Personal Relationships, 10*, 25-36. doi: 10.1111/1475-6811.00034
- Shackelford, T. K. (2001). Self-esteem in marriage. *Personality and Individual Differences, 30*, 371-390. doi: 10.1016/S0191-8869(00)00023-4
- Shakya, H. B. (2015). Affect and well-being similarity among older Indian spouses. *Aging and Mental Health, 19*, 325-334. doi: 10.1080/13607863.2014.933308
- Siegel, M. J., Bradley, E. H., Gallo, W. T., & Kasl, S. V. (2004). The effect of spousal mental and physical health on husbands' and wives' depressive symptoms, among older adults: Longitudinal evidence from the Health and Retirement Survey. *Journal of Aging and Health, 16*, 398-425. doi: 10.1177/0898264304264208
- Song, Z., Foo, M. D., & Uy, M. A. (2008). Mood spillover and crossover among dual-earner couples: A cell phone event sampling study. *Journal of Applied Psychology, 93*, 443-452. doi: 10.1037/0021-9010.93.2.443

- Sowislo, J. F., & Orth, U. (2013). Does low self-esteem predict depression and anxiety? A meta-analysis of longitudinal studies. *Psychological Bulletin, 139*, 213-240. doi: 10.1037/a0028931
- Steiger, A. E., Fend, H. A., & Allemand, M. (2015). Testing the vulnerability and scar models of self-esteem and depressive symptoms from adolescence to middle adulthood and across generations. *Developmental Psychology, 51*, 236-247. doi: 10.1037/a0038478
- Tambs, K., & Moum, T. (1992). No large convergence during marriage for health, lifestyle, and personality in a large sample of Norwegian spouses. *Journal of Marriage and the Family, 54*, 957-971. doi: 10.2307/353175
- Tower, R. B., & Kasl, S. V. (1996). Depressive symptoms across older spouses: Longitudinal influences. *Psychology and Aging, 11*, 683-697. doi: 10.1037/0882-7974.11.4.683
- Townsend, A. L. (2012). Innovative research methods for family gerontology. In R. Blieszner & V. H. Bedford (Eds.), *Handbook of families and aging* (pp. 61-88). Santa Barbara, CA: Praeger.
- Townsend, A. L., Miller, B., & Guo, S. (2001). Depressive symptomatology in middle-aged and older married couples: A dyadic analysis. *Journal of Gerontology: Social Sciences, 56B*, S352-S364. doi: 10.1093/geronb/56.6.S352
- Trzesniewski, K. H., Donnellan, M. B., & Robins, R. W. (2003). Stability of self-esteem across the life span. *Journal of Personality and Social Psychology, 84*, 205-220. doi: 10.1037/0022-3514.84.1.205
- Veit, C. T., & Ware, J. E. (1983). The structure of psychological distress and well-being in general populations. *Journal of Consulting and Clinical Psychology, 51*, 730-742. doi: 10.1037/0022-006X.51.5.730

- Wagner, J., Becker, M., Lüdtke, O., & Trautwein, U. (2015). The first partnership experience and personality development: A propensity score matching study in young adulthood. *Social Psychological and Personality Science*, *6*, 455-463. doi: 10.1177/1948550614566092
- Wagner, J., Lüdtke, O., & Trautwein, U. (2016). Self-esteem is mostly stable across young adulthood: Evidence from latent STARTS models. *Journal of Personality*, *84*, 523-535. doi: 10.1111/jopy.12178
- Wagner, J., Voelkle, M. C., Hoppmann, C. A., Luszcz, M. A., & Gerstorf, D. (2017). We are in this together: Dyadic patterns of self-esteem change in late-life couples. *International Journal of Behavioral Development*, Advance online publication. doi: 10.1177/0165025416679742
- Watson, D. (2004). Stability versus change, dependability versus error: Issues in the assessment of personality over time. *Journal of Research in Personality*, *38*, 319-350. doi: 10.1016/j.jrp.2004.03.001
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, *54*, 1063-1070. doi: 10.1037/0022-3514.54.6.1063
- Watson, D., & Humrichouse, J. (2006). Personality development in emerging adulthood: Integrating evidence from self-ratings and spouse ratings. *Journal of Personality and Social Psychology*, *91*, 959-974. doi: 10.1037/0022-3514.91.5.959
- Watson, D., Klohnen, E. C., Casillas, A., Simms, E. N., Haig, J., & Berry, D. S. (2004). Match makers and deal breakers: Analyses of assortative mating in newlywed couples. *Journal of Personality*, *72*, 1029-1068. doi: 10.1111/j.0022-3506.2004.00289.x

- Whiffen, V. E., & Aube, J. A. (1999). Personality, interpersonal context and depression in couples. *Journal of Social and Personal Relationships, 16*, 369-383. doi: 10.1177/0265407599163005
- Widaman, K. F. (2006). Missing data: What to do with or without them. *Monographs of the Society for Research in Child Development, 71*, 42-64. doi: 10.1111/j.1540-5834.2006.00404.x
- Williamson, H. C., Ju, X., Bradbury, T. N., Karney, B. R., Fang, X., & Liu, X. (2012). Communication behavior and relationship satisfaction among American and Chinese newlywed couples. *Journal of Family Psychology, 26*, 308-315. doi: 10.1037/a0027752
- Wothke, W. (2000). Longitudinal and multigroup modeling with missing data. In T. D. Little, K. U. Schnabel, & J. Baumert (Eds.), *Modeling longitudinal and multilevel data: Practical issues, applied approaches, and specific examples* (pp. 219-240). Mahwah, NJ: Erlbaum.

Footnotes

¹ In the studies by Segrin and colleagues, not all relationships examined were romantic relationships (i.e., partners who were dating, cohabiting, or married). More precisely, in Segrin et al. (2005), the sample included about 75% romantic relationships and 25% nonromantic relationships (e.g., close friends); in Segrin et al. (2007), the sample included about 77% romantic relationships and 23% nonromantic relationships; in Segrin et al. (2012), the sample included 83% romantic relationships and 17% nonromantic relationships. Nevertheless, the proportion of romantic relationships in these samples was sufficiently large to provide relevant background for the present research.

² All of the datasets have been used in previous publications. However, previous publications did not examine the research questions of the present article. There is one instance, in which the analyses overlap with analyses reported in a previous publication. Specifically, Schaffhuser et al. (2014) used data from the study Co-Development in Personality (CoDiP) to examine the relation between personality and relationship satisfaction in couples. In this context, they report on the prospective partner effect of self-esteem. However, this finding is not a central focus of Schaffhuser et al.'s article and, moreover, in the present research we use the CoDiP data to examine prospective partner effects of three additional constructs that were not examined by Schaffhuser et al. (i.e., life satisfaction, positive affect, and negative affect). We report the effect on self-esteem in the present article so that all information included in the meta-analytic aggregation across studies is readily available to readers.

³ As noted above, we did not use the first assessment of the LSG, which was conducted in 1971. At each of the following waves (i.e., the six waves included in the present research), all participants were eligible to participate, even if they did not provide data at one or more of the

earlier assessments. Therefore, it was possible that the number of participants, for whom data on study variables were available, increased from one wave to the next, as observed from the 1985 to 1994 assessments (i.e., Waves 1 to 4 of the present research).

⁴ For studies that included more than two waves, the models yielded more than one estimate for the longitudinal effects (i.e., one estimate for each time interval between waves). Although the coefficients were constrained to be equal across time intervals, the constraints were imposed on unstandardized coefficients (as typically recommended), which led to slight variation in the resulting standardized coefficients. Therefore, the standardized coefficients were averaged across time intervals.

Table 1

Descriptive Information on Studies

Study	<i>N</i> ^a	Age range at Wave 1	Number of waves	Time lag between waves
LSG	885	22–93 years	6	3 years
IYFP	451	29–68 years	4	1 year
LSDEC	300	22–49 years	3	1 year
CoDiP	236	16–92 years	2	2 years
MPI	186	18–61 years	2	6 months

Note. LSG = Longitudinal Study of Generations; IYFP = Iowa Youth and Families Project;

LSDEC = Longitudinal Study of Dual-Earner Couples; CoDiP = Co-Development in

Personality; MPI = My Partner and I.

^a Sample size used to compute the study weights for the meta-analytic computations, reflecting the number of couples who provided data on at least one of the study variables at one of the assessments.

Table 2

Means and Standard Deviations of the Measures (Averaged Across Waves) for Male and Female Partners

Variable	Male partner		Female partner	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
LSG				
Positive affect	0.77	0.27	0.77	0.26
Negative affect	0.28	0.29	0.30	0.30
Depression	1.45	0.40	1.51	0.43
Self-esteem	3.43	0.43	3.40	0.43
IYFP				
Positive affect	4.51	0.76	4.46	0.80
Depression	1.35	0.42	1.51	0.52
Self-esteem	4.00	0.49	3.99	0.56
LSDEC				
Depression	0.63	0.54	0.72	0.57
Self-esteem	3.44	0.46	3.42	0.49
CoDiP				
Life satisfaction	4.04	0.63	4.03	0.67
Positive affect	3.67	0.54	3.58	0.54
Negative affect	1.63	0.50	1.70	0.54
Self-esteem	3.42	0.43	3.28	0.50
MPI				
Life satisfaction	5.02	1.23	5.10	1.31
Depression	0.54	0.41	0.61	0.45
Self-esteem	4.18	0.69	4.06	0.80

Note. LSG = Longitudinal Study of Generations; IYFP = Iowa Youth and Families Project;

LSDEC = Longitudinal Study of Dual-Earner Couples; CoDiP = Co-Development in

Personality; MPI = My Partner and I.

Table 3

Test of Gender Differences in Actor and Partner Effects

Variable	df_A	df_B	Critical $\Delta\chi^2$	Observed $\Delta\chi^2$
LSG				
Positive affect	516	514	583.9	4.6
Negative affect	516	514	583.9	0.7
Depression	516	514	583.9	34.4
Self-esteem	354	352	413.8	0.2
IYFP				
Positive affect	222	220	150.8	0.7
Depression	222	220	150.8	4.1
Self-esteem	222	220	150.8	4.7
LSDEC				
Depression	120	118	65.9	0.6
Self-esteem	120	118	65.9	4.2
CoDiP				
Life satisfaction	48	46	29.6	1.9
Positive affect	48	46	29.6	0.1
Negative affect	48	46	29.6	2.0
Self-esteem	48	46	29.6	0.2
MPI				
Life satisfaction	48	46	25.4	0.7
Depression	48	46	25.4	0.4
Self-esteem	48	46	25.4	0.3

Note. Gender differences in actor and partner effects were tested by comparing the fit of two models, one that constrained the effects to be equal across male and female partners (Model A) and another that freely estimated the effects (Model B), using the test of small difference in fit (MacCallum et al., 2006). Given that the observed $\Delta\chi^2$ values were smaller than the critical $\Delta\chi^2$ values, the results indicated that equality constraints across male and female partners did not significantly worsen model fit. LSG = Longitudinal Study of Generations; IYFP = Iowa Youth and Families Project; LSDEC = Longitudinal Study of Dual-Earner Couples; CoDiP = Co-Development in Personality; MPI = My Partner and I.

Table 4

Fit of Models Tested

Variable	χ^2	<i>df</i>	CFI	TLI	RMSEA [90% CI]
LSG					
Positive affect	810.4*	516	.95	.94	.025 [.022, .029]
Negative affect	774.0*	516	.95	.94	.024 [.020, .027]
Depression	1024.2*	516	.96	.96	.033 [.030, .036]
Self-esteem	610.1*	354	.97	.97	.029 [.025, .032]
IYFP					
Positive affect	369.8*	222	.98	.98	.038 [.031, .045]
Depression	462.9*	222	.97	.97	.049 [.043, .055]
Self-esteem	306.0*	222	.99	.99	.029 [.020, .037]
LSDEC					
Depression	187.2*	120	.98	.98	.043 [.031, .055]
Self-esteem	135.9	120	1.00	.99	.021 [.000, .037]
CoDiP					
Life satisfaction	48.4	48	1.00	1.00	.006 [.000, .043]
Positive affect	74.6*	48	.98	.97	.048 [.025, .069]
Negative affect	58.4	48	.99	.99	.030 [.000, .055]
Self-esteem	36.7	48	1.00	1.00	.000 [.000, .021]
MPI					
Life satisfaction	81.1*	48	.98	.98	.061 [.037, .083]
Depression	65.5*	48	.99	.98	.044 [.005, .069]
Self-esteem	63.9	48	.99	.99	.042 [.000, .068]

Note. CFI = comparative fit index; TLI = Tucker–Lewis index; RMSEA = root-mean-square

error of approximation; CI = confidence interval; LSG = Longitudinal Study of Generations;

IYFP = Iowa Youth and Families Project; LSDEC = Longitudinal Study of Dual-Earner

Couples; CoDiP = Co-Development in Personality; MPI = My Partner and I.

* $p < .05$.

Table 5

Study-Level Findings on the Effects of Similarity, Stability, Mutual Influence, and Shared Environment

Variable	Similarity	Stability	Mutual influence	Shared environment
LSG				
Positive affect	.23*	.65*	.05*	.08*
Negative affect	.36*	.73*	.06*	.23*
Depression	.32*	.61*	.06*	.17*
Self-esteem	.15*	.79*	.01	.09*
IYFP				
Positive affect	.15*	.74*	.02	.12*
Depression	.08	.65*	.01	.17*
Self-esteem	.06	.84*	-.02	.14*
LSDEC				
Depression	.28*	.68*	.04	.01
Self-esteem	.12	.86*	.00	.14*
CoDiP				
Life satisfaction	.26*	.72*	.09	.32*
Positive affect	.24*	.70*	.18*	.04
Negative affect	.25*	.65*	.12*	.12
Self-esteem	.16*	.77*	.06	.07
MPI				
Life satisfaction	.36*	.79*	.16*	.04
Depression	.33*	.49*	.14*	.22*
Self-esteem	.15	.83*	.08*	.14

Note. The table shows standardized coefficients. LSG = Longitudinal Study of Generations;

IYFP = Iowa Youth and Families Project; LSDEC = Longitudinal Study of Dual-Earner

Couples; CoDiP = Co-Development in Personality; MPI = My Partner and I.

* $p < .05$.

Table 6

Meta-Analytic Aggregation of the Effects of Similarity, Stability, Mutual Influence, and Shared Environment

Effect	<i>k</i>	<i>N</i>	Weighted mean		Homogeneity (<i>Q</i>)
			effect size	95% CI	
Life satisfaction	2	422			
Similarity			.30*	[.215, .389]	1.26
Stability			.75*	[.708, .792]	2.75
Mutual influence			.12*	[.025, .214]	0.52
Shared environment			.20*	[.107, .291]	8.72*
Positive affect	3	1,572			
Similarity			.21*	[.161, .256]	2.34
Stability			.69*	[.658, .711]	9.33*
Mutual influence			.06*	[.012, .110]	4.27
Shared environment			.09*	[.036, .135]	1.06
Negative affect	2	1,121			
Similarity			.34*	[.285, .389]	2.72
Stability			.71*	[.685, .742]	4.34*
Mutual influence			.07*	[.014, .131]	0.67
Shared environment			.21*	[.151, .263]	2.38
Depression	4	1,822			
Similarity			.26*	[.214, .300]	20.48*
Stability			.62*	[.593, .650]	11.10*
Mutual influence			.05*	[.007, .098]	2.34
Shared environment			.15*	[.104, .194]	7.44
Self-esteem	5	2,058			
Similarity			.13*	[.084, .170]	2.91
Stability			.81*	[.800, .829]	17.84*
Mutual influence			.01	[-.029, .057]	1.89
Shared environment			.11*	[.068, .153]	1.60

Well-being (averaged across life satisfaction, positive affect, negative affect, and depression)	5	2,058			
Similarity			.26*	[.217, .298]	13.95*
Stability			.68*	[.654, .701]	1.25
Mutual influence			.06*	[.019, .105]	3.72
Shared environment			.13*	[.090, .176]	5.54
Difference well-being – self-esteem	5	2,058			
Δ Similarity			.13*	[.087, .172]	4.58
Δ Stability			-.13*	[-.177, -.092]	1.69
Δ Mutual influence			.05*	[.005, .091]	0.30
Δ Shared environment			.02	[-.021, .066]	10.50*

Note. Computations were made with a fixed-effects model. k = number of studies; CI = confidence interval.

* $p < .05$

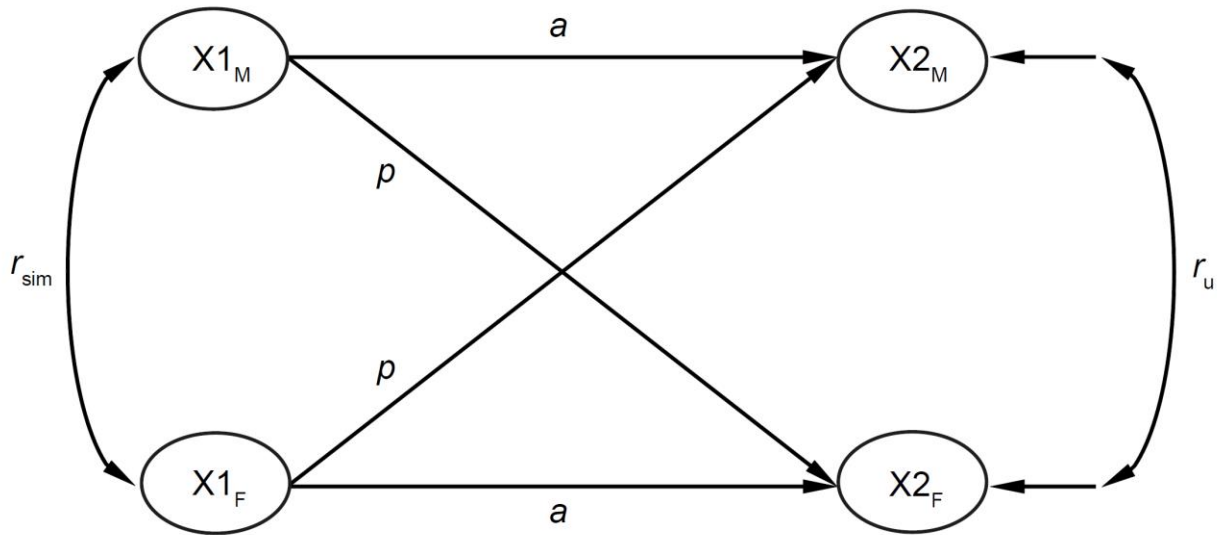


Figure 1. The figure illustrates the longitudinal actor-partner interdependence model. X denotes the construct measured in both partners at both Time 1 and Time 2. Subscripts denote whether the variable belongs to the male partner (M) or female partner (F). $X1$ = construct measured at Time 1; $X2$ = construct measured at Time 2; a = actor effect (i.e., stability effect); p = partner effect (i.e., mutual influence); r_{sim} = correlation between partners at Time 1 (i.e., similarity effect); r_u = correlation between the unexplained variances (i.e., effect of shared environment). The figure shows only latent constructs and omits observed variables.