UNIVERSITY of CALIFORNIA PRESS | Collabra: Psychology

Mu, W., et al. (2019). The Relationship between Self-Esteem and Depression when Controlling for Neuroticism. *Collabra: Psychology*, 5(1): 11. DOI: https://doi.org/10.1525/collabra.204

ORIGINAL RESEARCH REPORT

The Relationship between Self-Esteem and Depression when Controlling for Neuroticism

Wenting Mu[*,‡], Jing Luo[*,§], Sven Rieger[†], Ulrich Trautwein[†] and Brent W. Roberts[*,†]

Much research has examined the interplay of depression and self-esteem in an effort to determine whether depression causes self-esteem (scar model), or vice versa (vulnerability model). In the current longitudinal study (N = 2,318), we tested whether neuroticism served as a confounding variable that accounted for the association of depression and self-esteem, using both cross-lag models and latent growth models. We found neuroticism accounted for the majority of covariance between depression and self-esteem, to the degree that the scar and vulnerability models appear to be inadequate explanations for the relation between depression and self-esteem. Alternatively, neuroticism appears to be a viable cause of both depression and self-esteem and could explain prior work linking the two constructs over time.

Keywords: Depression; self-esteem; neuroticism; scar model; vulnerability model

Many theories of depression postulate that low self-esteem is a defining feature of depression (e.g., Beck, 1967; Brown & Harris, 1978). Indeed, low self-worth is one of the diagnostic criteria for depression in DSM-5 (APA, 2013). Empirically, a strong relationship has been observed between the two constructs. Cross-sectional correlations range from –.24 to –.79, depending on the samples and measures that were used (for a detailed review, see Orth, Robins, & Roberts, 2008).

Although the association between low self-esteem and depression has been well established, the nature of this relationship remains unclear. Two theories have been proposed to explain the relation between depression and self-esteem: the vulnerability model and the scar model. The vulnerability model proposes that low self-esteem is a critical causal factor that renders certain individuals to be more vulnerable to develop depression under stressful life events (e.g., Beck, 1967; Butler, Hokanson, & Flynn, 1994; Metalsky, Joiner, Hardin, & Abramson, 1993; Roberts & Monroe, 1992; Whisman & Kwon, 1993). Multiple pathways have been proposed that explain why people with lower self-esteem might be at higher risk for depression (see Orth et al., 2008 for a review). For example, according to Beck's (1967) cognitive theory of depression, negative beliefs about the self, which are central to low self-esteem, would contribute to the development of depressive disorders.

The scar model postulates that episodes of depression leave scars in the self-esteem system even after the remittance of a depression episode (Lewinsohn, Steinmetz, Larson, & Franklin, 1981; Rohde, Lewinsohn, & Seeley, 1990). According to this view, low self-esteem is a consequence of depression rather than a causal factor. Multiple pathways are assumed to underlie this relationship. For example, depression might diminish self-esteem by negatively altering the way in which individuals process self-relevant information, with those who have suffered depression being more likely to attend to, encode, and retrieve negative information about the self.

To date, the majority of research has provided evidence in favor of the vulnerability model. A growing body of longitudinal studies has found that low self-esteem prospectively predicts depression (e.g., Lewinsohn et al., 1988; Abela et al., 2006). The strongest evidence comes from studies that examined and compared both the prospective effect of self-esteem on depression and that of depression on self-esteem. For example, Orth, Robins and Roberts (2008) used two large longitudinal data sets and examined the relationships between low self-esteem and depression. Using cross-lagged panel models, they found low self-esteem predicted subsequent levels of depression, but depression did not predict subsequent levels of self-esteem, thus providing support for the vulnerability model, but not the scar model. This finding has been replicated in several subsequent studies (e.g., Orth, Robins, & Meier, 2009; Orth, Robins, Trzesniewski, Maes, & Schmitt, 2009; Rieger, Göllner, Trautwein, Roberts,

* Department of Psychology, University of Illinois, Champaign, IL, US

† Hector Research Institute of Education Sciences and Psychology, University of Tübingen, Tübingen, DE

‡ Department of Psychiatry, University of Pennsylvania, Philadelphia, PA, US

§ Department of Psychology, University of Southern California, Los Angeles, CA, US

Corresponding author: Wenting Mu (wenting.mu@pennmedicine.upenn.edu)

2016). However, it should be noted that a few studies have also examined the prospective effect of depression on self-esteem and found support for the scar model (Burwell & Shirk, 2006; Shahar & Davidson, 2003; Shahar & Henrich, 2010). In fact, a recent meta-analysis (Sowislo & Orth, 2013) yielded support for both the vulnerability model and the scar model, though the putative effect of self-esteem on depression was significantly stronger (double in size) than the effect of depression on self-esteem.

Neuroticism as a Confounding Variable

Whereas self-esteem has been considered a viable vulnerability factor for depression, there is another variable that may be a vulnerability factor for both self-esteem and depression, neuroticism. Although defined in various ways, the consensus definition is that neuroticism, at its core, is the propensity to experience negative emotions (Clark & Watson, 1999; Depue & Lenzenweger, 2001; Digman, 1997; Matthews et al., 2003; McCrae & Costa, 1997; Tellegen & Waller, 1997; Widiger, 2009). Neuroticism is believed to encompass many lower-order traits or facets. Most researchers agree anxiety-withdrawal, depression-unhappiness, vulnerability-stress reaction are facets of neuroticism; others would include angry hostility-aggression, impulsivity, inferiority, and dependency should be considered neuroticism as well (Ormel et al., 2013).

A strong relationship with neuroticism has been documented for self-esteem and depression in both cross-sectional and longitudinal studies. For example, Francis and James (1996) showed that emotional stability, a label commonly used for the low end of the neuroticism continuum, is associated with high self-esteem (r = .32) using Eysenck's two-dimensional model of personality and Rosenberg's self-esteem scale. Zeigler-Hill and colleagues (2015) found that self-esteem was positively associated with emotional stability using both self-reported scales (r = .45) and perceiver reported scales (r = .35). Also, in a large study of 326,641 participants, Robins et al. (2001) found that emotional stability was positively correlated with self-esteem (r = .50).

Neuroticism has also been shown to be a strong predictor for depression. For example, Kotov, Gamez, Schmidt, and Watson (2010) conducted a meta-analysis of 175 studies and found a strong positive association between neuroticism and depression (d = 1.33). Also, Kendler (2004) conducted a study on over 7500 twins trying to predict the onset of major depression and found an interaction effect between neuroticism and adversity. Individuals with high neuroticism were both at greater overall risk for major depression and were more sensitive to the depressogenic effects of adversity. The relationship between neuroticism and depression goes beyond cross-sectional level of analyses. It has been shown that change in neuroticism was also a predictor of change in depression, such that an increase in neuroticism was associated with an increase in depression (Chow & Roberts, 2014).

Despite significant covariances among the three constructs, they are essentially distinct constructs. For self-esteem and neuroticism, one particular useful way to look at their conceptual distinction is the perspective of core versus surface characteristics (Kandler, Zimmermann, & McAdams, 2014). Neuroticism is considered a core characteristic, which are largely consistent patterns of thoughts, feelings and actions across time and situations. By contrast, self-esteem is considered a surface characteristic, or characteristics that are believed to emerge much later, continue to evolve through lifespan, and are less stable or more environmentally malleable than core characteristics (McAdams & Pals, 2006). According to this point of view, it is believed that self-esteem is the by-product of the interaction of core characteristics, such as neuroticism and environmental influences (McCrae, 2009).

Empirical evidence on self-esteem and neuroticism seems to be in line with this view. First, compared to neuroticism, self-esteem has been found to be a less stable phenotype. Meta-analyses and longitudinal studies on representative samples have revealed that levels of rank-order stability are higher for neuroticism than self-esteem in adulthood (Trzesniewski, Donnellan, & Robins, 2003). Gene-environment interplay studies reveal that whereas neuroticism has been found to be more genetically based, self-esteem is found to be more subject to environmental influences, such as achievements, life stressors and failures (e.g., Kandler, Zimmermann, & McAdams, 2014). Heritability rates for neuroticism have been found to range from 40% to 60% (see Bartels & Boomsma, 2009, for a review), but found to range from 20% to 40% for self-esteem (e.g., Neiss, Sedikides, & Stevensen, 2002; Pedersen, Gatz, Plomin, Nesselroade, & McClearn, 1989). Therefore, it seems that self-esteem is a less stable and more environmentally malleable surface manifestation of personality.

Likewise, despite significant overlap, neuroticism and depression also differ from each other in fundamental ways. To begin with, while neuroticism is a trait like variable, depression is a state like variable. It is, by definition, a mental disorder, with onsets and episodes. Second, although depressivity, the predisposition to experience depression, is one facet of neuroticism, neuroticism is a much broader construct encompassing many more other facets, such as anxiety-withdrawal, vulnerability-stress reaction, hostility-anger (Ormel et al., 2013). Furthermore, neuroticism has been proposed to be an important vulnerability factor underlying the development of depression and this view has received considerable empirical support (see Klein, Kotov, & Bufferd, 2011 for a review). For example, studies show the rank-order stability of depression tends to be lower than that of neuroticism (Ormel et al., 2013). What is more, it should be noted the prospective association between neuroticism and depression has been well established even across long intervals of multiple years and adjusted for other psychiatric confounding variables. Last but not least, multiple longitudinal studies have found that the lifestyles of high-neuroticism individuals increase the likelihood of stressful experiences, and that these stressors in turn can trigger depression (Hankin, Stone, & Wright, 2010). Collectively, these findings strongly suggest neuroticism plays a very important role, at least partly,

in the complex pathways leading to the development of both low self-esteem and high depressive symptoms.

The Modeling Issue

The majority of work showing that low self-esteem prospectively predicts depression has relied on cross-lagged panel models. However, self-esteem and depression are conceptualized and measured as different types of variables. Global self-esteem, as it is usually studied in relationship to depression, is assessed and measured as a trait, denoting the "average tone of self-feeling" that each person carries around (Williams, 1995). It is a person's long-term, typical, affectively laden self-evaluation (Leary and Baumeister, 1995). By contrast, depression is an affective disorder that is considered to be episodic, with onset as well as remission. Most measures of depression are measures of states and not traits. Typical measures ask people to rate items based on how they have felt over the past two weeks (DSM-5, 2013). Given the different levels of conceptualization and measurement, it is not completely surprising to find the broader and more stable construct (i.e., self-esteem) often out-predicts the less stable construct (i.e., depression).

It is also questionable if cross-lagged panel models are the best way to establish the temporal order of longitudinal relationship among psychological constructs. It has recently been pointed out that standard cross-lagged panel models do not separate within- and between person effects and they assume that each person varies over time around the same mean (Hamaker et al., 2015). In other words, it is assumed there is no time invariant, trait-like individual differences that endure, an assumption that does not hold for most psychological variables (Fraley & Roberts, 2005). For example, it has been shown that there are ample individual differences in both the average level and change for major personality traits (Lüdtke et al., 2011), self-esteem (Wagner et al., 2013) as well as depression (Chow & Roberts, 2014).

Latent growth models are good tools to address this issue. These models separate the stable, between-person component and the within-person, changing aspect of any construct being examined. These models also can be used to estimate individual differences in both the initial levels of a variable (e.g., depression) as well as change over time (Hoffman, 2015). For this reason, latent growth models have been suggested to be useful in studying change in personality traits over time and determining the temporal orders of correlational relationships (Hamaker et al., 2015). They have also been used to study how individual differences in change in personality is related to change of other important outcomes, such as stress (Luo & Roberts, 2015) and mental health (Mu et al., 2016). As noted above, a positive relationship has been found between change in neuroticism and change in depression, such that increases in neuroticism were associated with increases in depression (Chow & Roberts, 2014). In another study, older men who were high in neuroticism at the beginning of the study and who increased in neuroticism over the course of the study experienced a higher risk of mortality than men who began the study low in neuroticism or men who decreased in neuroticism over time (Mroczek & Spiro, 2007). Thus, it is possible that growth in trait-like constructs, such as self-esteem, could be correlated with change in depression over and above antecedent standing (the level in a growth model) and that this is a better way of modeling the interplay of variables over time. Therefore, we will employ a variety of models so as to more thoroughly test the potential confounding of the relation between depression and self-esteem by neuroticism.

The Current Study

Given the aforementioned findings and links among self-esteem, depression, and neuroticism, the current study aimed to examine the relationship between self-esteem and depression while controlling for neuroticism. We employed a data set from a large longitudinal study tracking over 2000 German students in their early 20s that had been used in prior research to replicate the cross-lagged relation between self-esteem and depressive symptoms (Rieger et al., 2016). We investigated the relationship between self-esteem and depression while controlling for neuroticism using both cross-lagged panel models and latent growth models. We hypothesized that the relationship between self-esteem and depression would be reduced after controlling for neuroticism in both cross-sectional and longitudinal analyses.

Methods

The data come from a large, ongoing longitudinal German study (Transformation of the Secondary School System and Academic Careers; TOSCA; for a detailed overview see Trautwein, Neumann, Nagy, Lüdtke, & Maaz, 2010). The TOSCA study currently encompasses six time points. Data for self-esteem, depression and neuroticism are available for three waves. T1 is 2 years after graduation from high school (February to May, 2004). Participants completed an extensive questionnaire taking about 2 hours in exchange for a financial reward of 10 Euros. The second (T2) and third (T3) assessment took place from February to May, 2006 and from February to May, 2008, respectively. Again, participants completed an extensive questionnaire taking about 2 hours in exchange for a financial reward of 10 Euros.

Given this is a very large panel study that examined a very wide range of variables for an extended period of time, the current dataset has allowed many important questions regarding personality to be examined and research to be published. Among others, the most relevant ones have examined neuroticism (Lüdtke et al., 2011), self-esteem (Wagner et al., 2013), as well as self-esteem and depression (Rieger et al., 2016). However, it should be noted that the current analyses regarding self-esteem, depression and neuroticism have never been previously reported.

Participants

Data of the interested constructs (i.e., self-esteem, depression, and neuroticism) were available for n = 2,318 (64.0% female) individuals at T1, n = 1,912 (64% female) individuals at T2, n = 1,871 (63% female) individuals at T3. The sample size of the pooled data set is $N = 2{,}512$. Mean age

of participants was M = 21.5 years (SD = .8) at T1, M = 23.4 years (SD = .6) at T2, and M = 25.4 years (SD = .7) at T3.

For attrition analyses, we compared continuers, who completed all three time points, with dropouts, who participated only in the first wave. There were no significant differences on the study variables: self-esteem (d = –.00, p = .99), depression (d = –.01, p = .94), and neuroticism (d = –.01, p = .84). However, continuers were more likely to be female (OR = .76, p = .02), had better grade point averages (d = .20, p = .001) and performed better on a reasoning ability test (Heller & Perleth, 2000; d = 0.13, p = .03). Overall, the differences between continuers and dropouts were small ($|ds| \leq .20$).

Measures

Self-esteem. Self-esteem was measured at the trait level. The Rosenberg Self-esteem Scale (RSE; Rosenberg, 1965) was used to assess self-esteem: three items were administered: (a) "*At times, I think I am no good at all.*" (b) "*All in all, I am inclined to feel that I am a failure.*" and (c) "*I wish I could have more respect for myself.*" These items were translated into German. Participants were asked to rate these items using a likert-type 4-point scale ranging from 1 ("*not at all*") to 4 ("*totally*"). Internal consistency was good across all three waves (α = .84 at T1, .84 at T2, .86 at T3).

Depression. Depressive symptoms were assessed with the 15-item German version ("Allgemeine Depressionsskala"; ADS-K; Hautzinger & Bailer, 1993) of the Center for Epidemiologic Studies Depression Scale (CES-D; Radloff, 1977). A sample item was "*I felt lonely.*" Participants were asked to rate how often they have felt this way during the last week, using a 4-point likert-type scale (0 = "*rarely or none of the time*", 1 = "*sometimes*", 2 = "*frequently*", 3 = "*most of the time*"). Internal consistency was good across three waves (α = .90 at T1, .91 at T2, .91 at T3).

Neuroticism. Neuroticism was measured using the German version (Borkenau & Ostendorf, 1993) of the NEO Five-factor Inventory (NEO-FFI; Costa & McCrae, 1992). Extensive work on the German translation has demonstrated the instrument's high reliability, validity, and comparability with the English original (e.g., Borkenau & Ostendorf, 1993). A sample item was "*I often feel tense and jittery.*" The items were rated on a 4-point scale (1 = "*strongly disagree*", 4 = "*strongly agree*"). In-depth psychometric analyses of the 4-point response format show that this format has some advantages over a 5-point scale (Lüdtke, Trautwein, Nagy, & Koller, 2004). Internal consistency was good across three waves (α = .87 at T1, .88 at T2, .89 at T3).

Statistical Analyses

All models were estimated in the framework of longitudinal confirmatory factor analyses using Mplus 7.3 (Muthén & Muthén, 1998–2012). Two-sided statistical tests were performed at a level of significance of 5%. However, due to the observational character of our study, we rely on effect sizes and confidence intervals in addition to p-values (Groot, 2014).

The statistical procedure encompassed roughly three steps: First, to determine whether the three constructs should be modeled separately or as indicators of a common factor, we tested a series of models: 1) the one-factor model vs. the two-factor model of self-esteem and neuroticism; 2) the one-factor model vs. the two-factor model of depression and neuroticism; 3) the one-factor model vs. the two-factor model of self-esteem and depression; and 4) the one-factor model vs. the three-factor model of self-esteem, depression, and neuroticism. Second, to properly interpret latent variable change in longitudinal models, at least strong measurement invariance has to be established (Meredith, 1993; Meredith & Teresi, 2006). Thus, we specified a latent state model with imposed strong measurement invariance (same loadings and intercepts for each indicator over time) for all constructs within one model. This model served as our baseline model and we derived the means, standard deviations as well as latent correlations between all three constructs from it. Third, to investigate the prospective relationship between self-esteem and depression, we estimated a cross-lagged panel model (Model 1) and thereby reproduced the results from Rieger et al. (2016). Following this, we specified a cross-lagged panel model controlling for neuroticism at each time point (Model 2, see **Figure 1**). Fourth, to study interindividual difference in change over time we specified latent growth curve models. In a first step, we estimated three univariate latent growth models for each construct separately (Model 3a, 3b, 3c). Following this, we estimated a dual latent growth model for self-esteem and depression (Model 4). In a last step, we constructed a tri-variate latent growth model, to examine the relationship between self-esteem and depression while controlling for both the initial level as well as change in neuroticism over time (Model 5, see **Figure 2**).[1]

Missing data. To deal with missing values, we used full-information maximum likelihood estimation, as this procedure has been shown to produce less biased and more reliable results compared with the more conventional methods (e.g., listwise or pairwise deletion; Allison, 2003; Graham, 2009).

Model fit criteria and parceling strategy. Following the recommendations of Hu and Bentler (1998, 1999), we employed the comparative fit index (CFI), Tucker-Lewis-Index (TLI), and the root-mean-square error of approximation (RMSEA) to assess model fit. Self-esteem has three items and therefore was modeled on the item level. For each of the depression and neuroticism scales, we randomly aggregated the items into three parcels. Item parcels have been shown to produce more reliable latent variables than individual items when used as indicators (Little, Cunningham, Shahar, & Widaman, 2002; but see Marsh, Lüdtke, Nagengast, Morin & von Davier, 2013).

Results

Descriptive Statistics

Table 1 shows the means and standard deviations of the measures used in the current study. The latent correlations are depicted in **Table 2**.

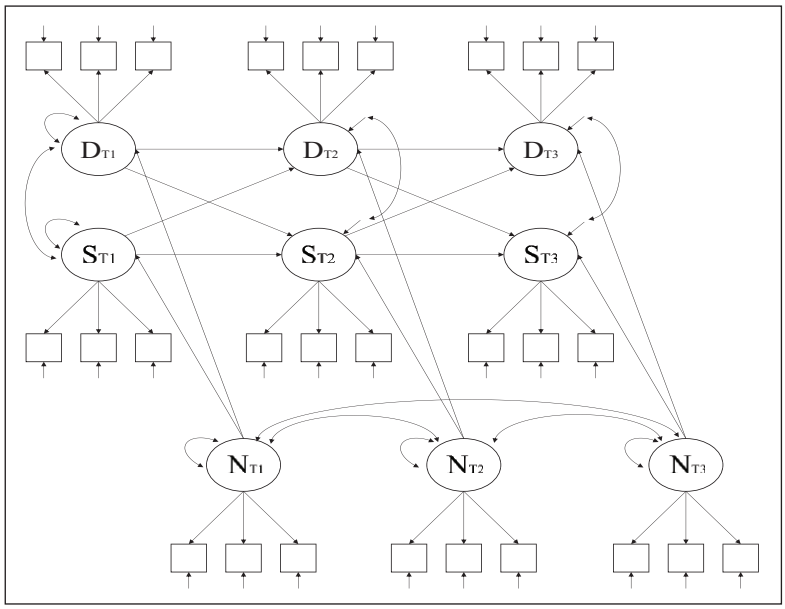

Figure 1: Cross-lagged panel model of self-esteem and depression controlling for neuroticism (Model 2) with longitudinal constraints on factor loadings and structural coefficients. D = Depression; S = Self-esteem; N = Neuroticism.

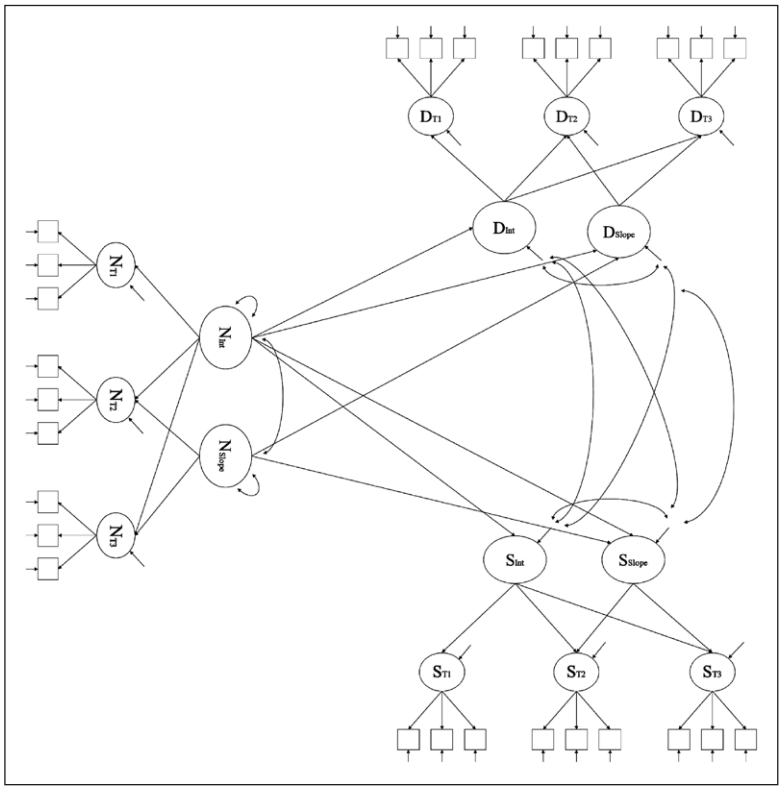

Figure 2: Latent growth curve model of self-esteem and depression controlling for neuroticism (Model 5). D = Depression; S = Self-esteem; N = Neuroticism.

Table 1: Means and Standard Deviations of Variables.

Variable	Time 1 (21 years)		Time 2 (23 years)		Time 3 (25 years)	
	M	*SD*	*M*	*SD*	*M*	*SD*
Self-esteem	3.42	.57	3.45	.58	3.56	.57
Depression	.66	.51	.60	.50	.56	.49
Neuroticism	2.17	.49	2.16	.52	2.11	.54

Note: *N* = 2,508. Means and SDs are extracted from Latent-state Models (Model 0).

Table 2: Latent Correlations among All Variables.

		1	2	3	4	5	6	7	8
1	Self-Esteem T1	–							
2	Self-Esteem T2	.73 [.69, .77]	–						
3	Self-Esteem T3	.63 [.59, .68]	.72 [.68, .76]	–					
4	Depression T1	–.61 [–.65, –.57]	–.48 [–.53, –.43]	–.41 [–.46, –.35]	–				
5	Depression T2	–.43 [–.48, –.38]	–.62 [–.66, –.58]	–.46 [–.52, –.41]	.47 [.42, .52]	–			
6	Depression T3	–.37 [–.42, –.31]	–.44 [–.49, –.38]	–.63 [–.68, –.59]	.39 [.33, .45]	.46 [.41, .52]	–		
7	Neuroticism T1	–.77 [–.80, –.74]	–.63 [–.67, –.60]	–.58 [–.62, –.54]	.64 [.61, .68]	.45 [.41, .50]	.42 [.38, .47]	–	
8	Neuroticism T2	–.65 [–.69, –.61]	–.81 [–.84, –.79]	–.66 [–.70, –.62]	.53 [.49, .57]	.65 [.62, .69]	.50 [.45, .54]	.78 [.76, .81]	–
9	Neuroticism T3	–.59 [–.64, –.55]	–.66 [–.70, –.62]	–.83 [–.86, –.81]	.46 [.42, .51]	.51 [.46, .56]	.68 [.65, .72]	.73 [.70, .76]	.78 [.76, .81]

Note: N = 2,508. Values in brackets are 95%-confidence intervals.

We first tested and compared a one-factor model vs. a two-factor model of self-esteem and neuroticism using data at Time 1 to determine whether these constructs were distinguishable. In the one-factor model, all indicators loaded on one common factor, separately for each time point; in the two-factor model, two latent factors (self-esteem and neuroticism) were specified. The two-factor model (χ^2 (8) = 45.20, CFI = .99, TLI = .99, RMSEA = .05, SRMR = .02) fit significantly better than the one-factor model (χ^2 (9) = 709.09, CFI = .87, TLI = .81, RMSEA = .19, SRMR = .06). Likewise, for depression and neuroticism, the two-factor model (χ^2 (8) = 83.14, CFI = .99, TLI = .98, RMSEA = .06, SRMR = .04) fit significantly better than the one-factor model (χ^2 (9) = 2085.30, CFI = .71, TLI = .52, RMSEA = .32, SRMR = .10). The same is true for self-esteem and depression, with the two-factor model (χ^2 (8) = 3.87, CFI = 1.00, TLI = .99, RMSEA = .04, SRMR = .01) fitting significantly better than the one-factor model (χ^2 (9) = 150.77, CFI = .78, TLI = .64, RMSEA = .26, SRMR = .11). Lastly, we also tested the one-factor model vs. the three-factor model of all the three constructs. As predicted, the three-factor model (χ^2 (24) = 6.31, CFI = .99, TLI = .98, RMSEA = .05, SRMR = .02) fit significantly better than the one-factor model (χ^2 (24) = 150.82, CFI = .78, TLI = .64, RMSEA = .26, SRMR = .11). Because the one-factor and the two/three-factor models are non-nested, we did not conduct formal test of difference in fit; however, the fit indices clearly indicate the favorability of the two-factor and the three-factor models. Similarly, previous research also found the two-factor model fit better for depression and self-esteem than the one-factor model (Orth et al., 2008). Thus, in the following analyses, the three constructs were modeled separately rather than as indicators of a common factor.

Results from Cross-lagged Panel Models

To answer our first research question (prospective relationship between self-esteem and depression), we first reproduced the results of Rieger et al. (2016) by constructing a regular cross-lagged panel model with freely structural coefficients (Model 1). In cross-lagged models, a latent variable at Time 2 is predicted by the same variable at Time 1 (the autoregressor) and the other latent variable at Time 1. The cross-lagged paths indicate the relation of one variable to the other, after controlling for the stability of the same variables over time (Finkel, 1995).

Model 1 showed a good fit to the data, χ^2 (122) = 387.35, CFI = .99, TLI = .98, RMSEA = .03, and SRMR = .04 (see also **Table 3**). Consistent with Rieger et al. (2016), the cross-lagged paths from self-esteem to depression were both statistically significant for T1 to T2 and T2 to T3 (β = –.22, SE = .04 and β = –.26, SE = .04, ps < .001). By contrast, the cross-lagged paths from depression to self-esteem were both nonsignificant (β = –.05, SE = .03, p = .09 and β = –.01, SE = .04, p = .76). The stability coefficients of self-esteem were β = .72, SE = .03 and β = .74, SE = .03 (ps < .001) and β = .34, SE = .04 and β = .32, SE = .04 (ps < .001) for depression.[2]

We next tested a second structural cross-lagged model (Model 2, see **Figure 1**), that is identical to model 1, except that we controlled for neuroticism at each time point. The model also showed a good fit to the data, χ^2 (301) = 1107.50, CFI = .98, TLI = .97, RMSEA = .03, and SRMR = .03 (see **Table 3**). The cross-sectional paths from neuroticism to self-esteem at the three time points were β = –.78, SE = .01, β = –.62, SE = .03, and β = –.66, SE = .03 (ps < .001); and the cross-sectional paths from neuroticism to depression were β = .66, SE = .02, β = .62, SE = .03 and β = .68, SE = .03 (all ps < .001), respectively. When controlling

Table 3: Fit Indices of the Models Tested.

Nr.	Model	N	Estimated Parameters	χ^2	*df*	CFI	TLI	RMSEA	SRMR	AIC	BIC
1	Model 1	2,506	67	387,35	122	.987	.983	.029	.035	45097,58	45487,95
2	Model 2	2,508	104	1107.50	301	.98	.97	.03	.03	59621.23	60227.27
3a	Model 3a	2,505	30	148,68	24	.983	.974	.046	.034	29704,10	29878,88
3b	Model 3b	2,499	30	68,57	24	.996	.994	.027	.019	17016,78	17191,49
3c	Model 3c	2,506	30	55,97	24	.997	.996	.023	.014	17959,95	18134,74
4	Model 4	2,506	67	354,56	122	.988	.985	.028	.026	45056,70	45447,07
5	Model 5	2,508	109	933,41	296	.982	.979	.029	.028	5942,82	60063,37

Note: CFI = comparative fit index; TLI = Tucker-Lewis-Index; RMSEA = root-mean-square error of approximation; SRMR = standardized root mean square residual; AIC = Akaike information criterion; BIC = Bayesian information criterion.

for neuroticism, the cross-lagged paths from self-esteem to depression switched from the negative direction to the positive direction. Self-esteem at T1 positively predicted depression at T2 (β = .12, *SE* = .04, *p* < .001) and self-esteem at T2 positively predicted depression at T3 (β = .16, *SE* = .03, *p* < .001). Likewise, the cross-lagged paths from depression to self-esteem were in the positive direction for both time points as well, with depression at T1 predicted self-esteem at T2 (β = .08, *SE* = .03, *p* < .01) and depression at T2 predicted self-esteem at T3 (β = .08, *SE* = .03, *p* = .004). The stability coefficients of neuroticism were *r* = .77, *SE* = .01 and *r* = .77, *SE* = .01 (*p*s < .001). Compared to Model 1 in which neuroticism was not controlled for, the stability coefficients of self-esteem dropped from .72 and .74 to .37 and .28 (*p*s < .001) and those of depression dropped from .34 and .32 to .21 and .21 (*p*s < .001). Given the high correlation of neuroticism with both self-esteem and depression, our interpretation of the switch from a negative relation in Model 1 (self-esteem to depression: –.22 & –.26; depression to self-esteem: –.05 & –.01) to a positive relation in Model 2 (self-esteem to depression: .12 & .16; depression to self-esteem: .08 & .08) is consistent with an artifact that results from controlling for all of the valid variance that explains the relation between self-esteem and depression.

Results from Latent Growth Models

Given the criticisms of cross-lagged regression models, we also tested the relations of all three variables using latent growth models. To assess the magnitude of interindividual differences in intraindividual change in neuroticism, self-esteem and depression, we constructed three univariate latent models with multiple indicators (Model 3a, 3b, 3c). All the three models showed excellent fit to the data (CFI and TLI > .95 and RMSEA and SRMR < .05; see **Table 3**). The latent intercepts represent the initial level of a personality trait at T1 and the variance of the latent intercepts indicates the amount of reliable individual differences at T1. The mean of the latent slope factors indicates the rate of change across the 2-year period and the variance of the slope indicates the amount of reliable individual differences in change. Addressing our primary question, the statistically significant variance components revealed that all three factors showed significant interindividual differences in both the initial level and the intraindividual change over time. Specifically, the variance of initial level is .20, *SE* = .01, *p* < .001 for neuroticism, .28, *SE* = .02, *p* < .001 for self-esteem, and .14, *SE* = .02, *p* < .001 for depression. Likewise, the variance of change was .02, *SE* = .01, *p* < .001 for neuroticism, .04, *SE* = .01, *p* < .001 for self-esteem, and .02, *SE* = .01, *p* = .007 for depression. In terms of the change direction, on average, neuroticism and depression declined over the 2-year period (*m* = –.03, *SE* = .01 and *m* = –.05, *SE* = .01, *p*s < .001), and self-esteem was found to increase across time (*m* = .07, *SE* = .01, *p* < .001). Taken together, all variables exhibited significant inter-individual difference in change over time.

Next, parallel to what we examined using cross-lagged panel models, we examined the relationship between self-esteem and depression using a dual latent growth model (Model 4). Model 4 showed a very good fit to the data, χ^2 (122) = 354.56, CFI = .99, TLI = .99, RMSEA = .03 and SRMR = .03. We examined the associations between self-esteem and depression by focusing on both the initial levels as well changes of the two constructs. Correlations among the latent intercepts reflect associations among the initial levels of the two variables at T1. We found the levels of self-esteem and depression were highly correlated (*r* = –.78, *SE* = .04, *p* < .001). Correlations among the latent slopes reflect associations between the changes of the two constructs across time. Like the pattern we observed regarding initial levels, change in self-esteem was found to be significantly negatively associated with change in depression (*r* = –.82, *SE* = .15, *p* < .001).[3]

Last, we constructed the same dual latent growth model but controlled for both the initial level as well as change in neuroticism over time (Model 5, **Figure 2**). Again, the model showed a good fit to the data (CFI and TLI > .95, RMSEA and SRMR < .05; see **Table 3**). The association between neuroticism and self-esteem or depression was extremely high: for initial level, β = –.83, *SE* = .02 and β = .81, *SE* = .03 respectively, *p*s < .001, and for change, β = –.82, SE = .08 and β = .81, *SE* = .12 respectively, *p*s < .001. When controlling for neuroticism, the association between initial levels of self-esteem and depression dropped from –.78 (Model 4) to –.35[4] (Model 5), *SE* = .10 (*p* = .001), and the

association between change in self-esteem and depression dropped from –.82 (Model4) to –.37 (Model5), *SE* = .49, *p* = .45. The magnitude of the relationship between the intercept of self-esteem and the slope of depression and vice versa was similar to that of Model 4 (*r* = .41, *SE* = .26, *p* = .12; *r* = .28, *SE* = .19, *p* = .16).

Discussion

The current research sought to address a set of fundamental questions regarding the relationship between self-esteem and depression. Two theories have been proposed to explain this relationship: the vulnerability model and the scar model. Although a growing body of research has supported the vulnerability model by finding a prospective relationship from low self-esteem to depression (e.g., Orth et al., 2008), two sets of observations raise further questions about this conclusion. First, another variable, neuroticism, has been shown to be strongly related to both self-esteem and depression. Such findings raise the possibility that the relationship between self-esteem and depression may be accounted for by their respective overlap with neuroticism. However, to date, no studies have explicitly tested this hypothesis. Second, the strongest evidence supporting the vulnerability model has come from research employing cross-lagged panel models. However, these cross-lagged panel models have recently been called into question and have been shown to provide biased estimates of the relation between variables like self-esteem and depression over time (Hamaker et al., 2015). To address the aforementioned problems in the past research, the present study examined the relationship between self-esteem and depression while controlling for neuroticism using a variety of modeling techniques. Specifically, we hypothesized that the relationship between self-esteem and depression would be significantly reduced after controlling for neuroticism.

We first sought to reproduce the basic findings of the cross-lag panel regression analyses and determine what effect controlling for neuroticism in these models would have. Like prior research with this sample, the cross-lagged panel regression analyses showed that self-esteem prospectively predicted depression and not the reverse when neuroticism was not incorporated into the model. When neuroticism was controlled for, the prospective relationships from self-esteem to depression or that from depression to self-esteem were not only reduced, but unexpectedly reversed to a significantly positive coefficient, suggesting the models may be mis-specified. In fact, Hamaker et al. (2015) showed that cross-lagged panel models sometimes may reveal reciprocal effects that do not exist. They further demonstrated that such problems often result from the inability of cross-lagged models to adequately separate the within-person and the between-person level when the constructs contain time-invariant, trait-like individual differences. Therefore, when using the original cross-lag panel model while controlling for neuroticism we went one step further than the methodological fix and specified the most likely confound. Consistent with this idea, controlling for the effect of neuroticism not only reduced the relationship between self-esteem and depression, but reversed it. The latter pattern most likely resulted from both the misspecification implicit in the cross-lagged panel model and the importance of neuroticism to the relation of self-esteem and depression.

To better estimate the static and dynamic relations between self-esteem, depression, and neuroticism over time, we modeled these variables using latent growth models. These models specify intercept and growth parameters, and can still be extended to include lagged relations from intercepts to growth parameters. Using these better specified models, we found that neuroticism accounted for most, if not all, of the association between both the level and change of self-esteem and depression. Specifically, in these growth models, the link between overall level in self-esteem and overall level in depression dropped by more than a half from –.60 to –.24 and the link between change in self-esteem and change in depression dropped from –.36 to –.01 when controlling for neuroticism. Meanwhile, both the level and change of neuroticism was highly correlated with those of self-esteem and depression. In addition, the stability coefficients of neuroticism were much higher (average .78) compared to those of self-esteem (average .33) or depression (average .21). Put together, these findings further suggest neuroticism serves as a confounding variable for both self-esteem and depression. That is neuroticism is a confound factor that, in part, explains the relation between the two variables and fully accounts for any dynamic relation between self-esteem and depression.

Our findings imply that neuroticism may be the cause of self-esteem and depression. These findings are also consistent with the abundance of evidence in the clinical literature, which shows that neuroticism predicts most forms of psychopathology, such as depression, anxiety, psychological distress, and substance abuse, to name a few (Kotov, Gamez, Schdmidt & Watson, 2010; Mu, Luo, Nickel & Roberts, 2016). The broad associations observed between neuroticism and various forms of psychopathology have not only led people to theorize neuroticism as a trait vulnerability factor underpinning the risk of developing many forms of psychiatric disorders (see a review for Klein, Kotov, Bufferd, 2011), but also as a higher order factor accounting for the high levels of diagnostic overlap and comorbidity among the wide range of psychopathology (Krueger & Markon, 2006; Watson, 2005). It appears that neuroticism also plays this role for the overlap between self-esteem and depression.

One finding of note was that even after controlling for neuroticism, the initial *levels* of self-esteem and depression were still significantly correlated (–.24). This suggests there is something left over between the self-esteem and depression, even after neuroticism was controlled for. One possibility is that there is something common to both self-esteem and depression, yet is not captured by neuroticism. This postulation is in line with the findings that refute the common factor model, in which self-esteem and depression are assumed to tap the same construct that overlaps highly with neuroticism (Orth et al., 2008). The inability of neuroticism to fully account for

the overlap between self-esteem and depression cautions concluding that self-esteem, depression or neuroticism are indistinguishable constructs.

Whereas the levels of self-esteem and depression is still significantly correlated controlling for neuroticism, the correlation coefficients between *changes* of self-esteem and depression dropped to almost zero when neuroticism was controlled for. One limitation of the growth modeling approach is that we did not measure these constructs often enough to get an optimal index of change. Thus, it is still possible that some small portion of self-esteem and depression are dynamically related over time. We suspect that better data will be needed to adequately test the relation among dynamic components of neuroticism, depression, and self-esteem. For example, more thorough and continuous assessments of self-esteem and depression, rather than assessment at several years' interval, would be necessary to provide more reliable estimates of change. Future research should endeavor to conduct deeper assessments of the constructs of interest more often to address such limitations.

What implications do these findings have for the vulnerability and scar models? Numerous studies and have been devoted to exploring this question, many of which involve very well-designed and rigorous longitudinal studies. Despite the accumulating evidence leaning towards the vulnerability model, our finding suggests that neither of these two models is adequate to address this question, at least for the age group examined in the current study (21–25), because they both omit an important confounding variable, neuroticism. Indeed, the relationship between self-esteem and depression disappeared or was even reversed when neuroticism was taken into account. Our finding suggests that future research should switch focus to the role of neuroticism in the development of self-esteem and depression. Does it represent some broad liability factor? What is the genetic or neural underpinning of this liability process?

Limitations and Future Directions

Some cautions regarding this study should be considered. We did not test our research question in any other datasets besides this one, which limits the generalizability of our results. For example, one limitation is the generalizability of results with participants from Germany to the United States or other cultures. Although there are certainly cultural differences between Germany and other countries, to date, no major differences have been documented on change of personality traits (Ludtke et al., 2011) or associations between self-esteem and depression (Rieger et al., 2016). Furthermore, research on cross-national comparisons have found the rates of depression are similar across countries (Weissman et al., 1996). Nevertheless, it is unclear to what degree cultural influences might affect the relations among neuroticism, self-esteem and depression. Future research should test and see if the findings of the current study can be generalized to other diverse samples.

Another caveat involves the generalizability of our results to other age groups. The TOSCA sample consists of students in young adulthood, a critical period in personality development marked by confluence of multiple developmental tasks (Arnett, 2000) and dramatic increase in multiple personality traits in putatively positive directions (Roberts, Walton & Viechtbauer, 2006). However, given that other age groups have shown differential change trajectories of personality traits (Roberts, Walton & Viechtbauer, 2006; Schwaba & Bleidorn, 2018) and depression (Hankin et al., 1998), the associations observed among neuroticism, self-esteem and depression could be affected by developmental challenges specific to other age groups. Future research should examine the associations among self-esteem, depression and neuroticism as well as their continuity and change in other life stages, such as adolescence or senior adulthood.

It should be noted that our measurement of personality traits and psychological functioning were all based on self-report data. Self-report measures reflect mostly the individual's own perspective of one's personality, behaviors and mood, and can be possibly confounded by individual differences in social desirability, response styles and level of insight. Future studies should assess personality, mood using other approaches and perspectives, such as observer ratings. Another issue of employing the same methodology (i.e., self-report) to assess all the three constructs is common method variances (Podsakoff et al., 2003), which could have inflated the observed associations among the constructs of interest. Indeed, in our studies, the absolute values of cross-sectional correlations among self-esteem, depression and neuroticism were quite high, ranging from .61 to .83. Future studies should employ multiple methods in measuring self-esteem, depression and neuroticism, to obtain more comprehensive estimates of associations among the three thus reducing the common method variance.

Relatedly, we measured self-esteem using three items from the original *Rosenberg Self-esteem Scale*, given the current project is part of large longitudinal panel study and it was difficult to include all items for each scale. Past research has shown that self-esteem can be measured adequately with only one item (Robins et al., 2001). Also, the internal consistency and test-retest reliability of the three-item version were similar to those of its full scale (Rosenberg, 1979). Nevertheless, the content validity of our three-item version has not been formally tested and future research should explore if our findings can be replicated when more thorough measurement of self-esteem is employed.

Conclusion

In conclusion, the present study significantly extends prior research on self-esteem and depression by controlling for neuroticism, and examining not only the concurrent, but also the dynamic relationships among the three variables. Our results suggest that neuroticism is a confound variable that, in part, explains the relation between self-esteem and depression and fully accounts for any dynamic relation between the two variables. It is clear from our results that the relationship between self-esteem and depression may not be meaningful when neuroticism is taken into consideration.

Data Accessibility Statement

All participant data and analysis scripts are available at the following link: https://osf.io/3rw7x/.

DOI: https://doi.org/10.17605/OSF.IO/3RW7X

Notes

[1] The study was not preregistered.

[2] Following Hamaker, Kuiper & Grasman's (2015) suggestions, we also fit the random-intercept cross-lagged panel model. We found, when the intercepts of the two constructs were explicitly modeled and controlled for, the prospective relationship from self-esteem to depression disappeared β = .004, p = .92 for T1 to T2, and β = .054, p = .26 for T2 to T3. However, the prospective relationship from depression to self-esteem became significant, β = –.07, p = .04 for T1 to T2, β = –.08, p = .04 for T2 to T3.

[3] The intercept of self-esteem was positively associated with the slope of depression (r = .37, SE = .086, p < .001). Given that depression declined over the 2-year period and the value of its slope was negative, the positive correlation suggests that higher initial levels of self-esteem was associated smaller decrease in depression over time. The intercept of depression was also positively associated with the slope of self-esteem (r = .25, SE = .09, p = .005). Given that self-esteem increased over the 2-year period and the value of its slope was positive, the positive correlation suggests that higher initial levels of depression is associated with greater increase in self-esteem over time.

[4] Coefficient is a residual correlation.

Funding Information

This work was supported by a grant to Ulrich Trautwein from the Ministry of Science, Research and the Arts of Baden-Württemberg (Az: 33-7532.20/735).

Competing Interests

The authors have no competing interests to declare.

Author Contributions

- Contributed to conception and design: Wenting Mu, Brent Roberts, Ulrich Trautwein
- Contributed to acquisition of data: Ulrich Trautwein
- Contributed to analysis and interpretation of data: Wenting Mu, Jing Luo, Sven Rieger, Brent Roberts
- Drafted and/or revised the article: Wenting Mu, Brent Roberts, Sven Rieger
- Approved the submitted version for publication: Brent Roberts, Ulrich Trautwein

Author Information

Sven Rieger was a doctoral student at the LEAD Graduate School & Research Network [GSC1028] at the time of the study, who was funded by the Excellence Initiative of the German federal and state governments.

References

Allison, P. D. (2003). Missing data techniques for structural equation modeling. *Journal of Abnormal Psychology*, *112*(4), 545. DOI: https://doi.org/10.1037/0021-843X.112.4.545

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (DSM-5□). *American Psychiatric Pub*.

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

Bartels, M., & **Boomsma, D. I.** (2009). Born to be happy? The etiology of subjective well-being. *Behavior genetics*, *39*(6), 605. DOI: https://doi.org/10.1007/s10519-009-9294-8

Baumeister, R. F., & **Leary, M. R.** (1995). The need to belong: desire for interpersonal attachments as a fundamental human motivation. *Psychological bulletin*, *117*(3), 497. DOI: https://doi.org/10.1037/0033-2909.117.3.497

Beck, A. T. (1967). *Depression: Clinical, experimental, and theoretical aspects*. University of Pennsylvania Press.

Borkenau, P., & **Ostendorf, F.** (1993). NEO Five Factor Inventory (NEO-FFI) by Costa and McCrae. Hogrefe, Göttingen.

Brown, G. W., & **Harris, T.** (1978). Social origins of depression: a reply. *Psychological Medicine*, *8*(4), 577–588. DOI: https://doi.org/10.1017/S0033291700018791

Burwell, R. A., & **Shirk, S. R.** (2006). Self Processes in adolescent depression: The role of self-worth contingencies. *Journal of Research on Adolescence*, *16*(3), 479–490. DOI: https://doi.org/10.1111/j.1532-7795.2006.00503.x

Butler, A. C., Hokanson, J. E., & **Flynn, H. A.** (1994). A comparison of self-esteem lability and low trait self-esteem as vulnerability factors for depression. *Journal of personality and social psychology*, *66*(1), 166. DOI: https://doi.org/10.1037/0022-3514.66.1.166

Chow, P. I., & **Roberts, B. W.** (2014). Examining the relationship between changes in personality and changes in depression. *Journal of Research in Personality*, *51*, 38–46. DOI: https://doi.org/10.1016/j.jrp.2014.04.007

Costa, P. T., & **McCrae, R. R.** (1992). Neo PI-R professional manual.

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

Francis, L. J., & **James, D. J.** (1996). The relationship between Rosenberg's construct of self-esteem and Eysenck's two-dimensional model of personality. *Personality and Individual Differences*, *21*(4), 483–488. DOI: https://doi.org/10.1016/0191-8869(96)0009515

Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual review of psychology*, *60*, 549–576. DOI: https://doi.org/10.1146/annurev.psych.58.110405.085530

Hamaker, E. L., Kuiper, R. M., & **Grasman, R. P.** (2015). A critique of the cross-lagged panel model. *Psychological*

methods, *20*(1), 102. DOI: https://doi.org/10.1037/a0038889

Hankin, B. L., Abramson, L. Y., Moffitt, T. E., Silva, P. A., McGee, R., & **Angell, K. E.** (1998). Development of depression from preadolescence to young adulthood: Emerging gender differences in a 10-year longitudinal study. *Journal of abnormal psychology*, *107*(1), 128. DOI: https://doi.org/10.1037/0021-843X.107.1.128

Hankin, B. L., Stone, L., & **Wright, P. A.** (2010). Corumination, interpersonal stress generation, and internalizing symptoms: Accumulating effects and transactional influences in a multiwave study of adolescents. *Development and Psychopathology*, *22*(1), 217–235. DOI: https://doi.org/10.1017/S0954579409990368

Hautzinger, M., & **Bailer, M.** (1993). Allgemeine Depressionsskala. [The German version of Center for Epidemiologic Studies Depression Scale.]. Göttingen: Hogrefe.

Hoffman, L. (2015). *Longitudinal analysis: Modeling within-person fluctuation and change*. Routledge.

Hu, L. T., & **Bentler, P. M.** (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological methods*, *3*(4), 424. DOI: https://doi.org/10.1037/1082-989X.3.4.424

Hu, L. T., & **Bentler, P. M.** (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling: a multidisciplinary journal*, *6*(1), 1–55. DOI: https://doi.org/10.1080/10705519909540118

Kandler, C., Zimmermann, J., & **McAdams, D. P.** (2014). Core and surface characteristics for the description and theory of personality differences and development. *European Journal of Personality*, *28*(3), 231–243. DOI: https://doi.org/10.1002/per.1952

Kendler, K. S., Kuhn, J., & **Prescott, C. A.** (2004). The interrelationship of neuroticism, sex, and stressful life events in the prediction of episodes of major depression. *American Journal of Psychiatry*, *161*(4), 631–636. DOI: https://doi.org/10.1176/appi.ajp.161.4.631

Klein, D. N., Kotov, R., & **Bufferd, S. J.** (2011). Personality and depression: Explanatory models and review of the evidence. *Annual review of clinical psychology*, *7*, 269–295. DOI: https://doi.org/10.1146/annurev-clinpsy-032210-104540

Kotov, R., Gamez, W., Schmidt, F., & **Watson, D.** (2010). Linking "big" personality traits to anxiety, depressive, and substance use disorders: A meta-analysis. *Psychological Bulletin*, *136*(5), 768–821. DOI: https://doi.org/10.1037/a0020327

Krueger, R. F., & **Markon, K. E.** (2006). Reinterpreting comorbidity: A model-based approach to understanding and classifying psychopathology. *Annual Review of Clinical Psychology*, *2*, 111–133. DOI: https://doi.org/10.1146/annurev.clinpsy.2.022305.095213

Lewinsohn, P. M., Hoberman, H. M., & **Rosenbaum, M.** (1988). A prospective study of risk factors for unipolar depression. *Journal of abnormal psychology*, *97*(3), 251. DOI: https://doi.org/10.1037/0021-843X.97.3.251

Lewinsohn, P. M., Steinmetz, J. L., Larson, D. W., & **Franklin, J.** (1981). Depression-related cognitions: Antecedent or consequence? *Journal of abnormal psychology*, *90*(3), 213. DOI: https://doi.org/10.1037/0021-843X.90.3.213

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*(2), 151–173. DOI: https://doi.org/10.1207/S15328007SEM0902_1

Lüdtke, O., Roberts, B. W., Trautwein, U., & **Nagy, G.** (2011). A random walk down university avenue: Life paths, life events, and personality trait change at the transition to university life. *Journal of personality and social psychology*, *101*(3), 620. DOI: https://doi.org/10.1037/a0023743

Ludtke, O., Trautwein, U., Nagy, G., & **Koller, O.** (2004). A validation of the NEO-FFI in a sample of young adults: Effects of the response format, factorial validity, and relations with indicators of academic achievement. *Diagnostica*, *50*(3), 134–144.

Luo, J., & **Roberts, B. W.** (2015). Concurrent and longitudinal relations among conscientiousness, stress, and self-perceived physical health. *Journal of Research in Personality*, *59*, 93–103. DOI: https://doi.org/10.1016/j.jrp.2015.10.004

Marsh, H. W., Lüdtke, O., Nagengast, B., Morin, A. J., & **Von Davier, M.** (2013). Why item parcels are (almost) never appropriate: Two wrongs do not make a right—Camouflaging misspecification with item parcels in CFA models. *Psychological methods*, *18*(3), 257. DOI: https://doi.org/10.1037/a0032773

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

McCrae, R. R. (2009). The five-factor model of personality traits: Consensus and controversy. In: Corr, P. J., & Matthews, G. (Eds.), *The Cambridge handbook of personality psychology*, 148–161. New York, NY, US: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511596544.012

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, *58*(4), 525–543. DOI: https://doi.org/10.1007/BF02294825

Meredith, W., & **Teresi, J. A.** (2006). An essay on measurement and factorial invariance. *Medical care*, *44*(11), S69–S77. DOI: https://doi.org/10.1097/01.mlr.0000245438.73837.89

Metalsky, G. I., Joiner, T. E., Hardin, T. S., & **Abramson, L. Y.** (1993). Depressive reactions to failure in a naturalistic setting: A test of the hopelessness and self-esteem theories of depression. *Journal of abnormal psychology*, *102*(1), 101. DOI: https://doi.org/10.1037/0021-843X.102.1.101

Mroczek, D. K., & **Spiro, A.** (2007). Personality change influences mortality in older men. *Psychological Science*, *18*(5), 371–376. DOI: https://doi.org/10.1111/j.1467-9280.2007.01907.x

Mu, W., Luo, J., Nickel, L., & **Roberts, B. W.** (2016). Generality or Specificity? Examining the Relation between Personality Traits and Mental Health Outcomes Using a Bivariate Bi-factor Latent Change Model. *European Journal of Personality*, *30*(5), 467–483. DOI: https://doi.org/10.1002/per.2052

Muthén, L. K., & **Muthén, B. O.** (1998–2012). Mplus user's guide, 7.

Neiss, M. B., Sedikides, C., & **Stevenson, J.** (2002). Self-esteem: A behavioral genetic perspective. *European Journal of Personality*, *16*(5), 351–367. DOI: https://doi.org/10.1002/per.456

Orth, U., Robins, R. W., & **Meier, L. L.** (2009). Disentangling the effects of low self-esteem and stressful events on depression: Findings from three longitudinal studies. *Journal of personality and social psychology*, *97*(2), 307. DOI: https://doi.org/10.1037/a0015645

Orth, U., Robins, R. W., & **Roberts, B. W.** (2008). Low self-esteem prospectively predicts depression in adolescence and young adulthood. *Journal of personality and social psychology*, *95*(3), 695. DOI: https://doi.org/10.1037/0022-3514.95.3.695

Orth, U., Robins, R. W., Trzesniewski, K. H., Maes, J., & **Schmitt, M.** (2009). Low self-esteem is a risk factor for depressive symptoms from young adulthood to old age. *Journal of abnormal psychology*, *118*(3), 472. DOI: https://doi.org/10.1037/a0015922

Ormel, J., Jeronimus, B. F., Kotov, R., Riese, H., Bos, E. H., Hankin, B., Oldehinkel, A. J., et al. (2013). Neuroticism and common mental disorders: Meaning and utility of a complex relationship. *Clinical psychology review*, *33*(5), 686–697. DOI: https://doi.org/10.1016/j.cpr.2013.04.003

Pedersen, N. L., Gatz, M., Plomin, R., Nesselroade, J. R., & **McClearn, G. E.** (1989). Individual differences in locus of control during the second half of the life span for identical and fraternal twins reared apart and reared together. *Journal of Gerontology*, *44*(4), P100–P105. DOI: https://doi.org/10.1093/geronj/44.4.P100

Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & **Podsakoff, N. P.** (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of applied psychology*, *88*(5), 879. DOI: https://doi.org/10.1037/0021-9010.88.5.879

Radloff, L. S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied psychological measurement*, *1*(3), 385–401. DOI: https://doi.org/10.1177/014662167700100306

Rieger, S., Göllner, R., Trautwein, U., & **Roberts, B. W.** (2016). Low self-esteem prospectively predicts depression in the transition to young adulthood: A replication of Orth, Robins, and Roberts (2008). *Journal of personality and social psychology*, *110*(1), e16. DOI: https://doi.org/10.1037/pspp0000037

Roberts, B. W., Walton, K. E., & **Viechtbauer, W.** (2006). Patterns of mean-level change in personality traits across the life course: A meta-analysis of longitudinal studies. *Psychological bulletin*, *132*(1), 1. DOI: https://doi.org/10.1037/0033-2909.132.1.1

Roberts, J. E., & **Monroe, S. M.** (1992). Vulnerable self-esteem and depressive symptoms: Prospective findings comparing three alternative conceptualizations. *Journal of Personality and Social Psychology*, *62*(5), 804. DOI: https://doi.org/10.1037/0022-3514.62.5.804

Robins, R. W., Tracy, J. L., Trzesniewski, K., Potter, J., & **Gosling, S. D.** (2001). Personality correlates of self-esteem. *Journal of Research in Personality*, *35*(4), 463–482. DOI: https://doi.org/10.1006/jrpe.2001.2324

Rohde, P., Lewinsohn, P. M., & **Seeley, J. R.** (1990). Are people changed by the experience of having an episode of depression? A further test of the scar hypothesis. *Journal of Abnormal Psychology*, *99*(3), 264. DOI: https://doi.org/10.1037/0021-843X.99.3.264

Schwaba, T., & **Bleidorn, W.** (2018). Individual differences in personality change across the adult life span. *Journal of personality*, *86*(3), 450–464. DOI: https://doi.org/10.1111/jopy.12327

Shahar, G., & **Davidson, L.** (2003). Depressive symptoms erode self-esteem in severe mental illness: A three-wave, cross-lagged study. *Journal of consulting and clinical psychology*, *71*(5), 890. DOI: https://doi.org/10.1037/0022-006X.71.5.890

Shahar, G., & **Henrich, C. C.** (2010). Do depressive symptoms erode self-esteem in early adolescence? *Self and Identity*, *9*(4), 403–415. DOI: https://doi.org/10.1080/15298860903286090

Sowislo, J. F., & **Orth, U.** (2013). Does low self-esteem predict depression and anxiety? A meta-analysis of longitudinal studies. *Psychological Bulletin*, *139*(1), 213–240. DOI: https://doi.org/10.1037/a0028931

Trautwein, U., Neumann, M., Nagy, G., Oliver, L., & **Maaz, K.** (Eds.) (2010). *Schulleistungen von Abiturienten: Die neu geordnete gymnasiale Oberstufe auf dem Prüfstand* [School achievement at the end of high school: Effects of the reform of upper secondary education]. Wiesbaden, Germany: VS Verlag für Sozialwissenschaften. DOI: https://doi.org/10.1007/978-3-531-92037-5

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: https://doi.org/10.1037/0022-3514.84.1.205

Wagner, J., Lüdtke, O., Jonkmann, K., & **Trautwein, U.** (2013). Cherish yourself: Longitudinal patterns and conditions of self-esteem change in the transition to young adulthood. *Journal of Personality and Social Psychology*, *104*(1), 148. DOI: https://doi.org/10.1037/a0029680

Watson, D. (2005). Rethinking the mood and anxiety disorders: A quantitative hierarchical model for DSM-V. *Journal of abnormal psychology*, *114*(4), 522. DOI: https://doi.org/10.1037/0021-843X.114.4.522

Weissman, M. M., Bland, R. C., Canino, G. J., Faravelli, C., Greenwald, S., Hwu, H. G., Lépine, J. P., et al. (1996). Cross-national epidemiology of major depression and bipolar disorder. *Jama*,

276(4), 293–299. DOI: https://doi.org/10.1001/jama.1996.03540040037030

Whisman, M. A., & **Kwon, P.** (1993). Life stress and dysphoria: The role of self-esteem and hopelessness. *Journal of Personality and Social Psychology*, *65*(5), 1054. DOI: https://doi.org/10.1037/0022-3514.65.5.1054

Zeigler-Hill, V., Holden, C. J., Enjaian, B., Southard, A. C., Besser, A., Li, H., & **Zhang, Q.** (2015). Self-esteem instability and personality: The connections between feelings of self-worth and the big five dimensions of personality. *Personality and Social Psychology Bulletin*, *41*(2), 183–198. DOI: https://doi.org/10.1177/0146167214559719

Peer review comments
The author(s) of this paper chose the Open Review option, and Streamlined Review option, and the peer review comments are available at: http://doi.org/10.1525/collabra.204.pr

How to cite this article: Mu, W., Luo, J., Rieger, S., Trautwein, U., & Roberts, B. W. (2019). The Relationship between Self-Esteem and Depression when Controlling for Neuroticism. *Collabra: Psychology*, 5(1): 11. DOI: https://doi.org/10.1525/collabra.204

Senior Editor: Simine Vazire

Editor: Simine Vazire

Submitted: 19 October 2018 **Accepted:** 25 January 2019 **Published:** 15 March 2019