Predictors of Individual Differences in Personality and Personality Change:

Findings From Cross-Sectional and Longitudinal Studies

Kumulative Dissertation

zur Erlangung des akademischen Grades

Doktor der Naturwissenschaften (Dr. rer. nat.)

im Fach Psychologie

an der Fakultät für Psychologie und Sportwissenschaft

der Universität Bielefeld

vorgelegt von

Julia Richter

Bielefeld, im Juli 2020

Erstgutachter: Prof. Dr. Rainer Riemann (Universität Bielefeld)

Zweitgutachterin: Prof. Dr. Anna E. Kornadt (Universität Luxemburg)

Betreuer: Prof. Dr. Christian Kandler (Universität Bremen) & Prof. Dr. Rainer Riemann (Universität Bielefeld)

Danksagung

Ich möchte mich bei allen Personen bedanken, die mich auf dem Weg zu und während der Arbeit an dieser Dissertation unterstützt haben. Ganz besonderer Dank gilt meinen Betreuern Prof. Dr. Christian Kandler und Prof. Dr. Rainer Riemann, die mir fachlich sowie persönlich viel Vertrauen und Freiraum geschenkt haben, und die es mir ermöglichten, meine Arbeit in Bielefeld sowohl zu beginnen als auch zu beenden. Herzlich danke ich auch meinen Koautor*innen für ihre inspirierenden und hochwertigen Beiträge, von denen alle Manuskripte fachlich sehr profitiert haben. Ich bedanke mich auch bei allen Studienteilnehmer*innen an SPeADy, pairfam und PEDES, ohne die unsere Forschung erst gar nicht möglich wäre.

Daneben danke ich meinen (ehemaligen) Kolleginnen für die unterstützende und motivierende Arbeitsatmosphäre. Der Austausch über die Freuden und Sorgen im wissenschaftlichen Alltag hat mir des Öfteren das nötige Durchhaltevermögen geschenkt. Auch meinen SPeADy-Kolleg*innen möchte ich für die längst vergangene tolle Zeit danken. Die gesamte Arbeitsatmosphäre der AE04 war sehr unterstützend, wozu alle Mitarbeiter*innen und nicht zuletzt Prof. Dr. Riemann intensiv beigetragen haben. Bei Prof. Dr. Anna E. Kornadt möchte ich mich darüber hinaus speziell für die Begutachtung meiner Arbeit bedanken. Anke Hufer-Thamm und Alexandra Zapko-Willmes danke ich darüber hinaus noch einmal ganz speziell für ihre Freundschaft im wissenschaftlichen Arbeitsfeld.

Danke außerdem meinen Korrekturleser*innen für ihre Sorgfalt, Geduld und Mühe, mir wertvolles Feedback zu dieser Synopse zu geben. Besonders möchte ich Alexandra Zapko-Willmes und Anke Hufer-Thamm für die wertvollen inhaltlichen sowie Soliman Keshta und Lisa Schumacher für die sprachlichen Anregungen danken.

Ich danke auch meiner Familie, die mir wohl neben einem privilegierten Umfeld das nötige Erbgut mitgegeben hat, um eine solche Aufgabe anzugehen (danke an Prof. Riemann für die mittäglichen Erinnerungen daran). Meinen Eltern Sabine und Bernd E. H. Richter danke ich vor allem, dass sie mich seit klein auf unterstützt haben, meinen Weg, und somit auch dieses Abenteuer, anzugehen. Meiner (zweieiigen) Zwillingsschwester Lisa danke ich für viele bereichernde Gespräche auf unserem (teilweise gemeinsamen) Weg durch’s Leben. Es ist wunderbar, eine Person zu kennen, die immer da ist!

Selbstverständlich gilt riesiger Dank meinem Partner Ingo Könemann. Du hast mich unermüdlich bestärkt und liebevoll unterstützt, dieses Projekt trotz aller Widrigkeiten abzuschließen. Und du hast jederzeit uneingeschränkt an mich geglaubt. Ich danke dir!

Content

List of Manuscripts

The dissertation contains the following manuscripts that are accepted by or submitted for publication in international peer-reviewed journals in the field of personality and differential psychology:

I. Richter, J., Zapko-Willmes, A., & Kandler, C. (submitted for publication in *Journal of Research in Personality*). **Age trends in explicit motives across the lifespan — Insights from self- and informant reports.**

II. Richter, J., & Finn, C. (submitted for publication in *PLoS ONE*). **Transactions between self-esteem and conflict in romantic relationships — A five-year longitudinal study.**

III. Richter, J., Zimmermann, J., Neyer, F. J., & Kandler, C. (2020). **Do sojourn effects on personality trait changes last? A five-year longitudinal study.** *European Journal of Personality.*

During the synopsis, I refer to these studies with the help of their Roman numerals. The manuscripts are appended after the end of this synopsis.

0. Summary

Personality is malleable across the whole lifespan. This phenomenon has been supported by multiple studies based on classical conceptions of personality traits (mainly the Big Five). While it is established that personality changes are at least in part rooted in naturally occurring (maturation) processes, personality is oftentimes not independent of the specific social contexts in which it unfolds. This notion has already been described in the last century via the paradigm of dynamic transactionism. Some related phenomena have drawn much attention since, resulting in theories like the social investment principle, which explains personality maturation by the investment in age-appropriate adult social roles. However, while there is growing evidence that (changes in) social environmental contexts seem to influence personality changes, there are still many unanswered questions in the debate on contextual predictors of personality and personality change.

While some of these contexts (e.g., positive relationship characteristics in romantic relationships) have been studied a lot, less is known about others (e.g., negative relationship characteristics). Moreover, it is unknown how long contextual influences on personality changes (e.g., in the face of sojourn experiences) last after the respective contexts have been left. The classical set-point theory states that most characteristics tend to return to an individual's particular level after the ending of certain experiences. However, this has barely been tested to date. In relation to all of these topics, there is a lot of missing information regarding broad personality characteristics, that is, personality conceptions beyond the Big Five.

My dissertation aims at widening the knowledge on individual differences in personality associated to specific natural contexts. To that end, I considered broad personality characteristics (i.e., explicit motives, self-esteem, and Big Five traits) and proxies of multiple natural contexts (i.e., age, relationship conflict, and student sojourns). I used data from three large datasets: The Study of Personality Architecture and Dynamics (SPeADy), the German Family Panel (pairfam), and the project Personality Development of Sojourners (PEDES). In addition to the perspective of predicting individual differences in personality levels and (long-term) personality changes, one aspect that all manuscripts share is the application of latent structural equation models (SEM). These have multiple advantages in comparison to manifest measurements – foremost the explicit assumption, modeling, and thus control of measurement error. Within the next paragraphs, I will shortly summarize my three studies.

Despite being relevant to shaping human goals across the whole lifespan, not much is known about age trends in explicit motives. According to lifespan theories, younger individuals strive more strongly for personal success and growth than older persons. In contrast, older individuals have fewer but closer relationships and stronger protection motives compared to younger individuals. The reason is said to lie in each individual's perceived time until death, which is directly associated to age. My first study (Richter, Zapko-Willmes, & Kandler, submitted for publication) examined this proposition by analyzing mean-level differences in explicit motives across a broad age range. My co-authors and I assessed explicit motives of individuals via self-reports and reports of up to two well-informed acquaintances, making a multi-rater approach possible. Another novelty in our study was the explicit inclusion of gender differences regarding age trends. First, we analyzed self-reports and informant reports separately to compare their concordance. Second, we combined them in a multivariate latent SEM to yield more accurate measures of the underlying constructs corrected for measurement error, rater-specific biases, and interrelations of constructs.

In a first step, we compared self- and informant ratings regarding their overall correlations and age-associations. Despite rater-specific differences in estimates (which seemed to be independent from the rater perspective), medium correlations of self-other agreement and inter-rater consensus across the whole sample indicated shared variance between raters that was not random. Splitting mean-levels per age in years, divergences in age trends between self- and informant reports mainly occurred for young males and elderly people, but were found to be largely independent of gender and relationship duration. With the latent approach, we aimed to control for these divergences.

Within the latent SEM, health protection and intimacy tended to increase, especially in the elderly, supporting the assumption of increased protection motivation in this age group. Moreover, we found lower levels of achievement, affiliation, personal growth, and power motives in older females than in younger females. While the same was true for males' affiliation and personal growth levels, males' power and achievement motives did not show any age associations. Moreover, the intimacy motive of middle-aged individuals showed a negative age association. While lifespan theories seem suitable to explain some of the age differences in levels of explicit motives, the findings outlined above led us to the assumption that windows of opportunity and cohort-specific differences in gender roles might also account for some of the findings. Our study is one of the first to indicate that well-informed acquaintances might provide incrementally valid information on targets' explicit motives.

Moreover, we showed that considering gender can lead to more nuanced findings, which might foster the revision of theories on age trends in motives.

Turning to a transactional approach, my second study (Richter & Finn, submitted for publication) was conducted to examine the longitudinal interplay between contextual and personality characteristics. Self-esteem is a personality characteristic that has been found to be associated with multiple contextual characteristics, especially romantic relationships. This can be explained by two contrasting perspectives: sociometer and self-broadcasting perspectives. Self-broadcasting perspectives view self-esteem as a prerequisite for establishing and maintaining positive social reactions in encounters. Sociometer perspectives, on the other hand, regard low self-esteem as an indicator of social exclusion, pointing towards failure regarding survival and reproduction. Both perspectives imply associations between self-esteem and romantic relationships regarding both positive and negative patterns. However, most previous studies have only focused on positive self-esteem–relationship dynamics. To understand partner dynamics more extensively, we therefore assessed negative dynamic processes in couples. To that end, we applied three bivariate actor–partner latent change models examining longitudinal associations between self-esteem and three aspects of perceived relationship conflict in stable heterosexual couples from both partners' perspectives. To yield the most parsimonious model, we equated effects across gender and time.

Besides initial correlations and correlated changes, we found that relatively higher perceived conflict frequency predicted relative decreases in self-esteem within individuals, above and beyond initial self-esteem levels. This implies the following: If one partner perceived and reported that there were more conflicts in their own relationship than in other couples' relationships, his or her self-esteem would also subsequently tend to decrease more (increase less) compared to all other participants. Moreover, we found small partner effects of dysfunctional (i.e., destructive, withdrawal) conflict styles on self-esteem changes. For example, the withdrawing conflict behavior of one partner as perceived and reported by the other negatively predicted changes in the first partner's self-esteem.

In contrast to positive relationship characteristics, which seem to be both predictive of and predicted by self-esteem, these results support sociometer perspectives on self-esteem in accordance with negative relationship characteristics. We concluded that negative relationship characteristics have a small, yet significant impact on self-esteem, but not vice versa. If this finding can be replicated, it might be a first hint that negative and positive dynamics in

romantic relationships do not necessarily function parallel to each other. However, we cannot fully preclude the possibility that there were self-broadcasting effects of self-esteem on perceived conflict, too, which we simply did not find in our study. It is noticeable that one partner's perception of the other partner's negative conflict styles impacted the other partner's self-esteem. This supports the notion that personality change occurs in a social context, that is, within social interactions.

In light of an increasing number of studies that show short-term or intermediate effects of student sojourns on Big Five trait changes, my third study (Richter, Zimmermann, Neyer, & Kandler, 2020) aimed at answering the question how sojourners' personalities develop in the long run. Three scenarios were conceivable: First, according to the classical set-point model, sojourn effects could turn back, resulting in reversing differences between sojourners and stayers (reversed effects). Second, differences between sojourners and stayers could accentuate in the long run, leading to increased differences (accentuated effects). And third, sojourn effects could lead to stable personality differences between sojourners and stayers (sustained effects). To find out which scenario was the most probable, we compared sojourners' long-term trait changes to their fellow students who stayed home using latent neighbor-change models across four time points (T1–T2–T3–T4). The assessments took place two weeks before the start of sojourners' stays abroad (T1), approximately six months after sojourn start (T2), approximately nine months after sojourn start (T3), and approximately five years after T1 (T4). In additional analyses, we differentiated between sojourn groups. While short-term sojourners had returned home at T3, long-term sojourners were still in their host country at that time.

Our results indicated positive short-term effects of a sojourn on changes in openness and agreeableness, and a negative effect on neuroticism change across T1–T2. Moreover, the univariate model implied positive effects on extraversion change across T1–T2 for short-term sojourners, while we did not find this effect in the multivariate model or for long-term sojourners. Additionally, the univariate model implied a positive sojourn effect on neuroticism change across T2–T3 for short-term sojourners, which was not found in the multivariate model. Across T3–T4, we did not find any sojourn effects. However, additional Wald tests and Bayesian model comparisons revealed that we could most probably rule out accentuated sojourn effects for agreeableness and neuroticism, and reversed sojourn effects for openness in the long run. Moreover, descriptive findings suggested that recurring sojourn

experiences may influence trait changes, leading to pronounced differences in openness and agreeableness between sojourners and their fellow students who stayed home.

These findings imply that different sojourn durations might have different effects on personality changes. For example, the positive T2–T3 effect on short-term sojourners' neuroticism implied an effect of the transition back home. However, diverging effects across sojourn groups were not robust across our analysis strategies. Regarding long-term personality changes, the patterns differed according to the trait under investigation. For openness, a set-point model of personality trait changes may thus not be applicable, while it might well be suitable for agreeableness and neuroticism. Besides the fact that personality traits differ in their short- and long-term susceptibility to a sojourn, it might be possible that a drastic social context change like a sojourn is able to slightly buffer normative decreases in openness in the long run.

Through these three studies, we gain knowledge by examining contextual predictors of individual differences in personality and personality changes. Although all reported effects were small, our findings suggest that (social) contexts are suited to predict (long-term) personality changes. In the long run, my findings might help gain a broader understanding of individuals' functioning in their social contexts. More specifically, by providing information on contextual predictors of differences in personality, my studies might help to conduct hypotheses on the reasons for these associations, and reveal information on underlying processes. For all of my studies, further research should, of course, first investigate more deeply how much of the found effects is indeed attributable to environmental factors and how these effects interact with expressions of biologically rooted selection and/or maturation processes.

I. Introduction

While it is evident even to laypeople that people differ from each other regarding their personality (Neyer & Asendorpf, 2018), less consensus has been reached among researchers regarding the question of how to conceptualize personality. A classical approach in the field of differential psychology measures personality via the Big Five traits (e.g., McCrae et al., 1999). This approach has the advantage of being very feasible and at least in part cross-culturally generalizable (McCrae et al., 2000). However, more recently several scientists have argued for the need of a broader conceptualization of the term personality (Kandler, Zimmermann, & McAdams, 2014; McAdams & Olson, 2010). Two broader conceptualizations, the *neo-socioanalytic model of personality* (Roberts & Wood, 2006) and the *actor–agent–author framework* (McAdams, 2013), have guided me through the conduction of my studies. I will explain them in the following section.

Besides debating on conceptualizing personality, researchers have for a long time elaborated on the questions of whether, how, and why personality changes (Roberts, Wood, & Caspi, 2008). While the *ontogenetic perspective* states that personality changes are mainly rooted in endogenous maturation processes (McCrae et al., 2000), the *environmentalist perspective* proposes the opposite position stating that an individual's characteristics develop mainly if not only because of environmental experiences (Watson, 1930). By increased methodological possibilities, inter alia, the assessment of and theories on personality changes have advanced to a more balanced view. More specifically, by now, it has been acknowledged that individual differences in personality do have a genetic basis, but personality is not independent of the environment (Bleidorn et al., 2019). Indeed, an individual's personality can especially manifest as adaptation to their social context (e.g., Denissen, van Aken, Penke, & Wood, 2013; Roberts & Jackson, 2008; Roberts, Wood, & Smith, 2005).

One study by Specht, Egloff, and Schmukle (2011), for example, reported age trends in the Big Five traits, which were partly attributable to social experiences in the form of major life events (e.g., marriage, separation, or child birth). Indeed, there is strong evidence that life events, such as forming a romantic relationship (e.g., Lehnart, Neyer, & Eccles, 2010) or experiencing a student sojourn (e.g., Zimmermann & Neyer, 2013), are associated with personality change. Bleidorn and colleagues (2013) also reported strong universal personality change patterns across age. Moreover, the authors found differences in age-related personality change across cultures. This study shows that while age is often seen as reflecting intrinsic

maturation, it might also serve as a proxy for the cultural background. More specifically, while there seems to be an underlying intrinsic maturation across age, the exact timing of particular personality change processes also seems to be interwoven with socio-cultural factors. A similar argumentation applies to gender. It is obvious that social gender mainly depends on biologically rooted differences, which are proven to be associated with differences in personality (e.g., Schmitt, Realo, Voracek, & Allik, 2008). For example, Graham et al. (2020) recently reported slightly higher mean levels in neuroticism for females than for males. Gender differences manifest even and especially in countries that try to provide egalitarian environments (Schmitt et al., 2008). While the obvious conclusion from these findings is that gender differences in personality reflect biological differences (cf. Costa, Terracciano, & McCrae, 2001), differences in the social context are aligned to these differences (Merkel, 2020). The confounding of biological and societal factors makes it especially interesting to compare gender effects across different cohorts (study I).

The aforementioned studies are promising ways to pursue the prediction of personality and personality change. And while there is growing evidence on the predictive value of (changes in) social environmental contexts, there are also still many open questions. This especially applies to personality conceptions beyond the Big Five and specific contextual factors (e.g., negative relationship characteristics). Besides, we know almost nothing about possible long-term influences of social contexts after they have been left. With this dissertation, my goal is to answer some of these questions. To that end, I examined individual differences in personality characteristics and changes in association with age, gender, and social contextual variables (i.e., relationship conflict, student sojourns). I mainly refer to the predictors of personality as proxies of contextual predictors in my dissertation, although this is of course not the only way to characterize them. Before describing the three research questions of the present dissertation in more detail, I define how I approached personality and personality change in this work.

II. Personality and Personality Change

With the aim of investigating predictors of individual differences in personality and personality change, we must first consider how to define personality. Importantly, from the perspective of a differential psychologist, personality is usually described as an individual's characteristic pattern of feeling, thinking, and behaving in comparison to a certain reference population (Neyer & Asendorpf, 2018). This way, it is not only possible to compare

individuals within one sample regarding their personality *levels*, but also to detect covariates (e.g., predictors) of these level differences. A classical approach to personality is the five-factor model (McCrae & Costa, 1999). Here, personality is characterized by the Big Five *traits*, that is, openness to experience, conscientiousness, extraversion, agreeableness, and neuroticism. In this conception of personality, the authors differentiate between biologically rooted traits and characteristic adaptations, also reflecting environmental influences. My approach throughout this dissertation was based on more modern approaches to personality. While classical approaches usually equate traits with personality, I will only refer to traits in the form of the Big Five.

a. Personality beyond the Big Five

Although traits are a very useful tool to describe major differences between individuals and have been used for decades to do so, personality is more than the Big Five (McAdams & Olson, 2010; Roberts & Wood, 2006). Lately, differential psychologists have argued for more extensive frameworks to characterize personality that include all relatively enduring tendencies of typical feelings, attitudes, strivings, and behavioral tendencies (Kandler et al., 2014). Two theoretical models that conceptualize personality more broadly than the five-factor model are the *neo-socioanalytic model of personality* (Roberts & Wood, 2006) and the *actor–agent–author framework* (McAdams, 2013). While both models approach traits, motives, and narratives at the same analysis level, the actor–agent–author framework additionally takes self-esteem into consideration. These frameworks guided me through the conduction of the three studies. More specifically, in my work, I aimed to apply a broader view on *personality characteristics* by exemplarily employing *explicit motives*, *self-esteem*, and the *Big Five traits*.

These characteristics can be arranged at different layers. According to the actor–agent–author framework, personality traits belong to the first layer of a person – the *social actor* (McAdams, 2013). They define how an individual sees themselves or is seen by others in terms of typical cognitive, emotional, and behavioral patterns (McAdams & Olson, 2010). Motives belong to the second layer – the *motivated agent* –, and can be seen as the driving forces of action planning (McAdams, 2013). More specifically, motives underlie and shape specific goals and plans. An individual's self-esteem in this framework might function as a monitor to evaluate whether previous and present behaviors are directed towards or hindering an individual from important life goals (McAdams & Olson, 2010). If an individual suffers from low self-esteem, this could indicate their failure to accomplish some of these life goals.

While the actor–agent–author framework sees the reaching of general life goals as predictors of self-esteem, the *sociometer perspective* applies a similar view regarding social relationships (Leary, Tambor, Terdal, & Downs, 1995). From this perspective, low self-esteem indicates failure to be adequately socially integrated, while high self-esteem emerges from social inclusion. The *autobiographical author* represents the third layer in the actor–agent–author framework (McAdams, 2013), which refers to life narratives that individuals create to give their experiences meaning (McAdams & Olson, 2010). In my work, I focused on the first and second layer of a person.

b. Assessing personality change

Having established how I approached personality in my work, I am now switching to the question what *personality change*[1] is. Statistically, change is oftentimes conceptualized in contrast to continuity (see, for example, Roberts & Mroczek, 2008), and can be described in relative and absolute terms, respectively (Roberts et al., 2008). Researchers in the field of personality psychology are usually interested in universal patterns of personality and personality change across or within a population (Specht et al., 2014). To that aim, they observe and report *normative*, average personality change across a certain sample (Denissen et al., 2011) and *individual differences* in these average change patterns (Roberts et al., 2008; Specht et al., 2014).

Normative change is captured via *mean-level change* in a characteristic of interest *across* one sample (Specht et al., 2014). It is measured by absolute differences in the mean levels of a personality characteristic between two time points (compare Caspi & Roberts, 2001), thereby indicating mean increases or decreases (Roberts et al., 2008). Studying mean-level change across a sample gives insight into more universal patterns, which is important to yield a background against which to study individual differences (Denissen et al., 2011). Individual differences in personality change are usually measured via three approaches: *rank-order consistency*, *variance* around the mean change, and *mean-level differences* in change. Rank-order consistency is the relative ordering of individuals *within* the sample. It is usually reported as a test–retest correlation indicating the stability of the relative placement of an

[1] In the literature, the term personality *development* often occurs instead of change (e.g., Denissen, van Aken, & Roberts, 2011). Although development inherently implies change, it can easily be understood as a normatively evaluated direction of the latter. For example, young adults' personality traits conscientiousness, agreeableness, and neuroticism change in a socially desirable way in direction of more mature behavior (e.g., Bleidorn, Kandler, Riemann, Spinath, & Angleitner, 2009; Roberts, Walton, & Viechtbauer, 2006), which might be termed as developmental progress. Throughout this work, I tried to avoid expressions that might imply an evaluation, even of normative change patterns. Thus, I will mainly use the term *change*(*s*).

individual regarding a characteristic within their reference group (see Roberts & DelVecchio, 2000). Another, more recent way to capture within-person variability is via the variance around the average change index (Specht et al., 2014). For the purpose of identifying predictors of personality change, it is moreover fruitful to consider how mean-level differences in personality change are associated with these predictors (Roberts et al., 2008; Steyer, Schmitt, & Eid, 2000). For example, individuals who experience a specific life event might show higher increases or decreases in a personality characteristic in comparison to individuals who have not experienced it. It is important to note that this pattern can emerge independently from the average change in a personality characteristic in a given population. This kind of pattern could be identified by associating the life event to mean-level differences in personality change.

Assessing personality change requires multiple-wave data, usually captured over a long period of time (e.g., multiple yearly assessments). Besides the difficulty of recruiting participants and maintaining samples during *longitudinal studies*, analyzing data of multiple assessments brings specific challenges. For example, structural continuity, also referred to as *measurement invariance*, is an important prerequisite to be able to model and interpret personality change in terms of content (i.e., the persistence of correlational patterns over time or across groups). Besides many other advantages, *latent structural equation models* (SEMs) allow us to model and test restrictions regarding measurement invariance. This makes them especially valuable for longitudinal models that require structural continuity to meaningfully interpret personality changes. Moreover, the modeling of personality constructs as *latent variables* allows us to explicitly model measurement error (Steyer, Mayer, Geiser, & Cole, 2015), which leads to more accurate assessments in terms of reliability, and eventually to higher measurement accuracy in comparison to manifest modeling approaches. In many SEM software, missing data can be imputed by applying the full information maximum likelihood (FIML) estimation technique (Enders, 2011). As the FIML estimation uses all available information to impute missing values, it is especially useful for longitudinal studies and large survey studies where drop-out (across measurement occasions or within one long assessment session) is often inevitable. In all three studies of my dissertation, I applied latent SEMs.

One type of SEM are *latent change models* (LCM; McArdle, 2009). In these models, *true* change between two measurement occasions is modelled as their difference (see **Figure 1**; Steyer, Eid & Schwenkmezger, 1997). As latent difference variables portray intraindividual change (Steyer et al., 1997), it is possible to investigate correlates (e.g., predictors) of individual differences in intraindividual personality change (Steyer et al., 2000). This is why latent difference variable are sometimes termed *true intraindividual change scores* (Steyer et al., 1997). In comparison to the closely related latent growth curve models, LCM are less restrictive and easier to apply for more than two measurement occasions (Steyer et al., 1997). In two of my studies, I applied LCM to capture true intraindividual change in association with contextual covariates.

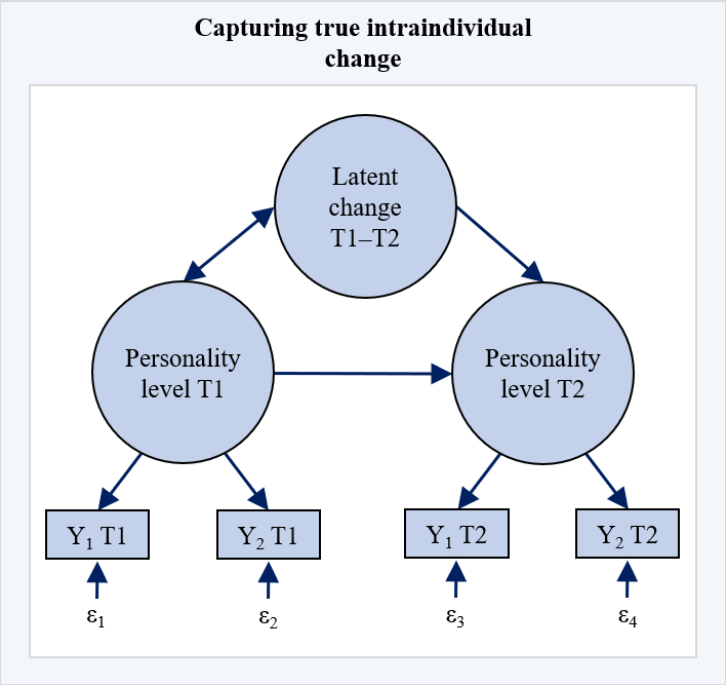

Figure 1. Classical conceptualization of true intraindividual change between two measurement occasions (adapted from study III). The underlying latent personality level (circle) is assessed via two manifest observations (rectangles) at each time point (T1 and T2), respectively, to ensure identification of the model (Steyer et al., 1997). Measurement error (ε) is allowed to vary across time.

c. Previous findings on personality change

Shifting the focus to findings concerning personality change, it has been established that personality characteristics show both continuity and change across the lifespan (Roberts & Wood, 2006). More specifically, rank-order continuity in traits increases at least until age 50 (*cumulative continuity principle*; Roberts & DelVecchio, 2000), while the possibility of personality change still persists into adulthood (Caspi & Roberts, 2001). The possibility of environmental influences on personality at any age has been acknowledged by the *plasticity*

principle (Roberts & Wood, 2006; Roberts et al., 2008). More specifically, the authors define a person as an open system, which inevitably interacts with the social environment. Importantly, according to the authors, a person's thinking, feeling, and behaving are susceptible to environmental influences during the whole life.

Importantly, behavioral patterns do not occur in a vacuum, but within socially relevant experiences (McAdams, 2013). People select themselves into environments that are consistent with their personality, usually referred to as *selection effects* (e.g., Denissen, Ulferts, Lüdtke, Muck, & Gerstorf, 2014; Denissen, Luhmann, Chung, & Bleidorn, 2019). According to Bleidorn and colleagues (2014, 2019), trying to identify environmental predictors of individual differences in personality might inform us how personality unfolds in a social context (see also Roberts & Wood, 2006). Controlling for selection effects helps to observe, for example, how the *average* personality is affected by certain environmental characteristics. Using this kind of approach, there is growing evidence that changes in social environments (foremost life events) are small, yet significant predictors of individual differences in personality changes above and beyond selection effects (e.g., Jackson, Thoemmes, Jonkmann, Lüdtke, & Trautwein, 2012; Lehnart et al., 2010; Lüdtke, Roberts, Trautwein, & Nagy, 2011; Specht et al., 2011), which is termed *socialization effects* (e.g., Zimmermann & Neyer, 2013). In studies II and III, I examined socialization effects after controlling for initial differences in personality. It should be noted that both the kind of environment and the kind of personality characteristic under investigation can have an impact on the direction and magnitude that is associated with personality changes (Bleidorn et al., 2012), which I have addressed in study III.

d. Theories on personality change

There are several theories aiming to explain why socialization effects can shape naturally occurring personality changes, and why these are usually small. The *social investment principle*, for example, states that age-graded social roles can foster personality changes as adaptation processes in direction of greater maturity (Helson, Kwan, John, & Jones, 2002; Roberts et al., 2005). According to the *self-regulation theory* (Denissen et al., 2013), personality changes are adaptation processes to social norms, which depend on an individual's regulatory capacities. While both frameworks are in line with the findings of normatively increasing mean levels in agreeableness, conscientiousness, and emotional stability in young adulthood (*maturity principle*; Roberts & Wood, 2006), the self-regulation theory also offers explanations to maturity decreases in accordance with new social roles. For

example, experiencing childbirth has sometimes been found to be accompanied by following decreases in conscientiousness (Denissen et al., 2019).

According to the actor–agent–author framework (McAdams & Olson, 2010), the individual is constantly trying to balance social roles, skills, and plans. This way, personality change across the lifespan can be explained by changing adaptation processes following changing demands, while personality continuity might be due to stable narratives. The neo-socioanalytic framework moreover addresses both the reasons for occurring and non-occurring personality changes in connection with social environmental factors. For example, some personality consistency might be explained by the increasing development, commitment, and maintenance of consistent social roles (*role continuity principle*), and thus one's identity (*identity continuity principle*; Roberts & Wood, 2006). The *sociogenomic personality psychology* explains this process as follows: Daily behavior unfolds and is associated with the environment an individual lives in. If behavior is shown consistently over a long period of time, it manifests as personality (Roberts & Jackson, 2008). Following from this, enduring behavioral changes can eventually lead to lasting changes in people's thinking, feeling, and behavior – in short, their personality (Roberts, 2006). In the course of my dissertation, I will refer back to these theories, integrating my findings on predictors of personality levels and changes into a broader context.

III. Findings on Predictors of Individual Differences in Personality and Personality Change

My dissertation was dedicated to the overall question which predictors shape individual differences in personality and personality change. To capture personality more broadly than traditionally, I used different personality characteristics in all three studies. For two of the studies, I used ratings of close acquaintances to rate personality or behavioral characteristics, thus controlling for rater-specific biases. In study I, one wave of data was accessible to investigate age trends in explicit motives across a broad age range. Thus, I approached age trends in explicit motives from a normative lifespan perspective. Focusing on individual differences in personality change at the population level, study II and study III examined social environmental predictors of individual differences in personality change. More specifically, in study II, I was interested in self-esteem–conflict transactions in stable romantic couples, while study III aimed at investigating possible long-term effects of a

sojourn on personality trait changes; both studies capturing a time span of five years. To that end, I applied LCM.

In the following sections, I will present the research questions I addressed, give an overview over the methods I used, and summarize the findings from my studies against the backdrop of prior research. After that, I will integrate my findings into a more global context. More specifically, I will both address answers and open questions that emerge from my findings against the backdrop of prior work and theories. I will close with an outlook on possible future work to more extensively capture predictors of individual differences in personality and personality change.

a. Study I: Tracking age trends in motives across the lifespan by combining self- with informant reports

From the perspective of a psychological self as social actor, motivated agent, and autobiographical author, motives are highly relevant to drive an individual's goals and behavior over the life course (McAdams, 2013). According to this framework, an individual develops a sense of being a motivated agent by the end of childhood. From that time on, personal plans and projects are said to become a central feature of the self. The assumption that explicit motives underlie specific life goals seems to be common consensus. More specifically, while motives reflect more complex personal strivings towards multiple underlying outcomes, for example intimacy, achievement, or status (Bleidorn, Kandler, Hülsheger, Riemann, Angleitner, & Spinath, 2010), specific life goals are oriented towards short- or intermediate term outcomes (Roberts & Robins, 2000), for example, graduation, having children, or living in a lakeside villa (study I). To date, there are multiple studies examining specific life goals, mostly using open frequency ratings (e.g., Ebner, Freund, & Baltes, 2006). However, empirical evidence on age trends in the more general, underlying motives is scarce (as a recent exception, see the study by Bühler, Weidmann, Nikitin, & Grob, 2019). My first study addressed this gap.

We tested the assumption of lifespan theories that motives shift from personal success and growth towards fewer but closer relationships and protection (Brandstädter, Rothermund, Kranz, & Kühn, 2010; Carstensen, 2006). More specifically, we hypothesized that protective motives (i.e., health protection and intimacy) show higher mean levels in older individuals compared to younger individuals, and expansive motives (i.e., achievement, affiliation, personal growth, power) show lower mean levels in older individuals compared to younger

individuals. We used data from the first wave of the German Study of Personality Architecture and Dynamics (SPeADy; Kandler, Penner, Richter, & Zapko-Willmes, 2019) comprising self- and informant reports of participants between 14 and 94 years of age.

Usually, personality characteristics are examined via self-rating questionnaires. This is mainly because researchers assume that self-raters have an acceptable introspection and because this approach is highly economic. Combining self-ratings with ratings of well-informed acquaintances yields additional information on self-other agreement (Neyer & Asendorpf, 2018). Usually, in the combination of self- and other rating, it is assumed that the self-rating is valid, and the higher the correlation between self- and informant rating, the better the observability of the personality characteristic (Neyer & Asendorpf, 2018). This conception does not always need to be accurate. Close friends, family, or partners might have access to observable behavior that is not necessarily conscious to the self-rater, for example, due to self-deception tendencies (Wagerman & Funder, 2007). This way, informant ratings can be an incrementally valid source of information beyond self-reports (Kim, Di Domenico, & Connelly, 2018). To infer this conclusion, it is necessary to have more than one informant rating. This way, it can be tested whether inter-rater consensus (i.e., correlations between two informants) is substantially higher than zero, that is, non-random (study I). In my first study, we first compared consensus between self-reports and reports of up to two well-informed acquaintances, respectively, across all participants.

Moreover, we applied a multivariate latent SEM using the composite scores of both rater perspectives to benefit from the advantage that SEMs allow to control for rater-specific biases (Dobewall & Aavik, 2016; Funder & Sneed, 1993). This way, we obtained the most valid ratings in terms of convergent validity. Please see **Figure 2** for an overview of the measurement model. SEMs also allow to model and examine differences in multiple groups at the same time. Unlike previous studies in this field, we included gender as a predictor to examine main gender effects as well as gender–age interactions on mean levels in motives (study I). Not only could we control for imbalances in sample sizes and associated biases in estimates this way, but we could also test and establish measurement invariance across genders, and were thus able to compare effect sizes of age trends in motives for females and males.

Self-other agreement and inter-rater consensus were of medium size and comparable to each other. This is important as it shows that a common knowledge between observers exists that is probably not shared by the self-rater (cf. Kandler et al., 2010). In total, correlations indicated shared variance despite rater-specific error. Moreover, highest deviations between the mean ratings of both perspectives (i.e., self- and informant rating) were received in men and in the elderly. Although this indicated that the self-rater's gender and the relationship duration between self- and informant rater might be moderators on self-other agreement, more extensive analyses did not support this notion for most motives. Thus, despite small moderator effects, shared variance between rater perspectives suggested that a latent modeling of the underlying motives would increase reliability of estimates. In general, this also implied that observers seem to have access to information relevant to judge explicit motives. This is a highly important finding showing the possible benefit of informant ratings on this matter, which might help increase their use in future motivational research.

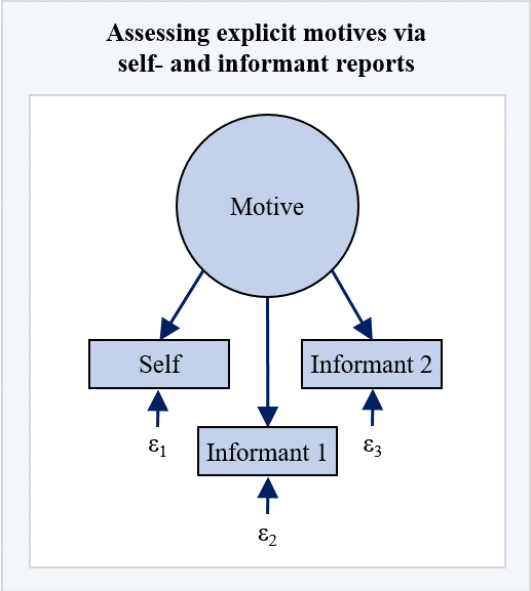

Figure 2. The underlying latent motives (circle) is assessed via up to three manifest observations (rectangles). Measurement error (ε) is allowed to vary across raters. Adapted from study I.

Results of the multivariate SEM including six motives yielded support for increased importance of protective motives (i.e., health protection and intimacy) in the elderly as compared to younger individuals. Moreover, as predicted, expansive motives affiliation and personal growth showed negative associations with age. However, not all results were as we had expected: Achievement and power showed negative age associations in females, but not in males. Moreover, health protection did not show any age associations for individuals younger than 50 years, and intimacy followed negative age trends in midlife, speaking against our prediction between the ages 30 to 45 years.

Our findings partly supported lifespan theories predicting that motives shift from achievement, affiliation, personal growth, and power towards intimacy and health protection. However, the pattern of results was more complex than these theories predict; especially for younger ages and for men. This led us to assume that windows of opportunity are important to explain age differences in motives, too (Heckhausen, 1997). Moreover, these findings show that gender is an important predictor of motives across the lifespan. We explained some of the findings by contextual, partly gender-specific factors, such as societally ascribed normative roles (e.g., Helson et al., 2002). However, biological differences might be better suited to explain personality differences across gender in cultures that try to provide an egalitarian environment (Schmitt et al., 2008), even though gender equality is still not perfectly achieved to date (see Merkel, 2020, on the circumstances in Germany). As frameworks like the sociogenomic personality psychology suggest, societally formed normative gender roles and zeitgeist might be interwoven with biologically driven gender and age differences (Roberts & Jackson, 2008).

b. Study II: Investigating self-esteem–conflict transactions in romantic relationships from a dyadic perspective

With the aim to shed light on possible predictors of individual differences in personality changes, in my second study, I turned to a longitudinal approach. In accordance with previous findings on self-esteem and positive relationship characteristics (e.g., Erol & Orth, 2014; Mund et al., 2015), we examined transactions between self-esteem and perceived conflict in romantic relationships. From the perspective of the social investment principle (e.g., Helson et al., 2002), it is obvious to assume that engaging in a partnership accelerates personality changes in direction of greater maturation. Several studies have supported this notion, also with regard to stronger increases in self-esteem (e.g., Lehnart et al., 2010). According to the actor–agent–author framework, this might be explained by reaching important life goals (McAdams, 2013), while the sociometer perspective explicitly sees social inclusion as prerequisite for self-esteem (Leary et al., 1995). Other studies found that personality seems to be a better predictor of relationship characteristics than vice versa (e.g., Neyer & Asendorpf, 2001). With regard to self-esteem, this can be explained by the *self-broadcasting perspective* (Srivastava & Beer, 2005): The proposition that an individual's self-evaluation predicts their behavior in social situations, which in turn evokes the expected positive or negative reactions in their encounters.

From a paradigm of *dynamic transactions* (Magnusson, 1990), both personality and the environment are medium-term constant, but susceptible to long-term changes. Changes in one or the other are based on processes within the respective domain. However, importantly, the person influences the environment and vice versa. To date, personality–relationship transactions in romantic relationships have mostly been reported for forming a relationship in general (Lehnart et al., 2010; Wagner, Lüdtke, Jonkmann, & Trautwein, 2013) and regarding positive relationship aspects, such as relationship satisfaction (Mund, Finn, Hagemeyer, Zimmermann, & Neyer, 2015). Although it is known that relationships comprise multiple characteristics (Fletcher, Simpson, & Thomas, 2000) of positive and negative evaluative valence, there is a surprising lack of studies investigating negative relationship characteristics (study II).

To understand person–situation dynamics more profoundly, we thus wanted to shed light on negative personality–relationship transactions. One relationship feature of negative evaluative character is perceived relationship conflict (Peterson Tilden & Galyen, 1987). Against the backdrop of the aforementioned theories and in accordance with previous findings on self-esteem and positive relationship characteristics (e.g., Erol & Orth, 2014; Mund et al., 2015), we examined whether self-esteem and perceived conflict are negatively intertwined across time, both within and across partners. The data was obtained from the German Family Panel (pairfam; Brüderl et al., 2018; for a more detailed description, see Huinink, Brüderl, Nauck, Walper, Castiglioni, & Feldhaus, 2011). While self-esteem and perceived conflict frequency were self-rated, each participant reported their perception of their partner's dysfunctional conflict styles, captured as destructive behavior and withdrawal.

We applied three adapted longitudinal dyadic actor–partner interdependence models (Kenny, Kashy, & Cook, 2006) following the approach by Mund and colleagues (2015) to portray both partners' perspectives in couples' complex personality–relationship transactions. This way, we examined selection and socialization effects of both partners' perspectives in stable romantic couples at the same time and under control of the respective other across 5 time points (see **Figure 3**). Thus, self-esteem–conflict transactions were not only captured within both partners, i.e., *within-person effects* (here called *actor effects*). Moreover, including both partners in the gender-controlled LCM allowed us to test interpersonal dynamics, i.e., *between-person effects*, between self-esteem and perceived conflict (here called *partner effects*).

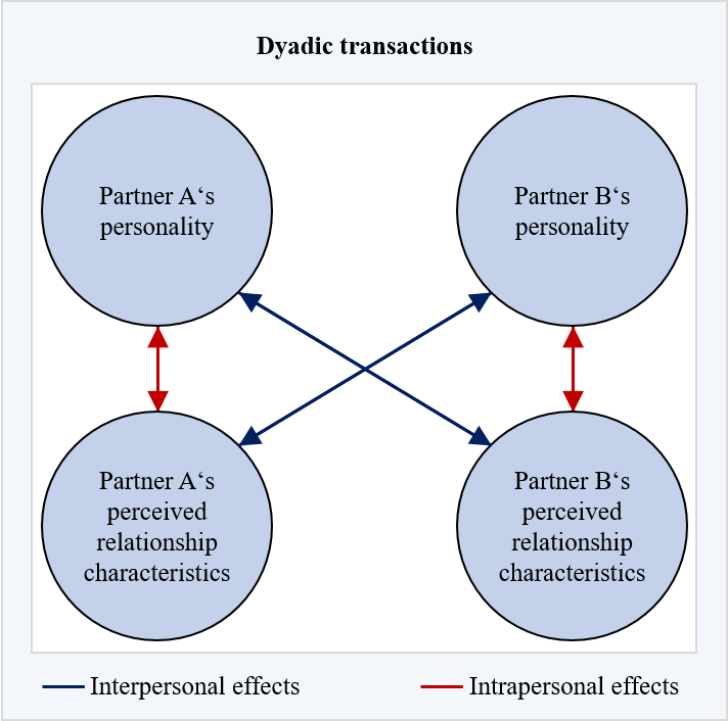

Figure 3. Assessing personality–relationship interdependencies (simplified model, adapted from Mund, Finn, Hagemeyer, & Neyer, 2016).

Self-esteem negatively correlated with perceived conflict frequency and styles both within persons and between partners. Moreover, we found moderate rank-order consistencies of all variables, indicating between-person variability in change trajectories. Level-change effects of perceived conflict frequency and style were able to explain a small part of this variability. More specifically, we found an actor effect of perceived conflict frequency on subsequent self-esteem change beyond and above prior self-esteem levels. Actor effects of perceived conflict styles on self-esteem changes tended to show the same pattern, but were only significant at $p < .10$. Partner effects of perceived destructive behavior and withdrawal on self-esteem changes above and beyond prior self-esteem levels indicated the following, here described using the couple Annie and Dan: The more Dan perceived destructive or

withdrawal behavior in Annie, the steeper Annie's self-esteem consequently decreased. Additionally, we found two positive change-change effects on the within-person level. That is, increases in perceived conflict predicted subsequent increases in self-esteem beyond and above initial self-esteem levels, and vice versa. These findings supported the notion that self-esteem and perceived conflict are longitudinally intertwined, and might be a hint on ceiling or floor effects.

Although all level-change effects were very small, they yielded support for perspectives that see self-esteem as being a monitor of social inclusion instead of being a predictor (e.g., Leary et al., 1995). We found one negative actor effect of perceived conflict frequency on self-esteem change. Put differently, the more conflict Dan perceived at a particular time point, the more his self-esteem decreased or the less it increased compared to the total sample. Moreover, we found that perceived conflict styles had a between-person effect on self-esteem changes, indicating mutual negative dynamics between partners. As described earlier, this means that the amount of Annie's destructive or withdrawal behavior reported by Dan predicted subsequent changes in Annie's self-esteem. Put differently, the more destructive or withdrawal behavior Dan reported about Annie, the more her self-esteem subsequently decreased (or the less it increased) in comparison to the total sample. We attribute this finding to Dan's lower recognition of Annie's dysfunctional behavior during conflicts, leading her to subsequently feel socially rejected. Based on these findings, we concluded that both higher perceived conflict frequency and lower recognition in conflict situations might function as indicators of social rejection, thereby negatively predicting subsequent self-esteem changes (Leary et al., 1995).

Importantly, if one partner perceived negative conflict styles in their relationship, this impacted the other partner's self-esteem. From a health care perspective, this finding is relevant if we consider the predictive effect of self-esteem on physical and mental health (cf. Rieger, Göllner, Trautwein, & Roberts, 2008; Trzesniewski, Donnellan, Moffitt, Robins, Poulton, & Caspi, 2006). Moreover, it stresses the notion that personality change occurs in a social context (e.g., Roberts & Wood, 2006). For couple consultants, not only is it important to know that both partners suffer from perceived conflict by decreasing self-esteem, it is also exciting to see that it might be possible to positively influence both partners' self-esteem via decreasing their perceived conflict. Importantly, this effect can of course be mediated by other behavioral variables. For example, Annie's withdrawing behavior in conflict situations might

increase the risk of Dan being unfaithful. In sum, however, the change of conflict behavior in the first place might still be a good starting point for interventions.

On a more general level, it seems possible that negative dynamics within romantic relationships do not function parallel to positive dynamics. Moreover, our findings imply that relationship effects on personality exist, and might even be independent of personality effects on relationship characteristics in stable romantic relationships. If this conclusion is true, it partly supports a contextualist perspective (Lewis, 2001), implying that personality changes are rooted in contextual influences. However, it is probable that we did not find very small but true self-broadcasting effects on perceived conflict changes. See, for example, diverging results in the meta-analysis by Harris and Orth (2019), supporting a more balanced mutual dynamic between self-esteem and the social context. At this point, we cannot rule out either of these possibilities.

c. Study III: Examining the sustainability of personality trait changes in response to student sojourns

In my second study, we found that very small social environmental effects on personality changes do exist. Asking the question how enduring effects of social environments on personality changes are, in my third study, I turned from a transactional approach to a more static assessment of the social environment, namely in the form of student sojourns. Student sojourns have been established to promote personality maturation in young adulthood across short and intermediate time intervals beyond and above selection effects (Greischel, Noack, & Neyer, 2016; Niehoff, Petersdotter, & Freund, 2017; Zimmermann & Neyer, 2013). Although this is an important finding and can help to deepen the knowledge on personality change processes in accordance with life events, employers in the German economy are more interested in how sojourners personality changes in the long run (DAAD, 2015). From a theory-building perspective, this question is not trivial either (study III). However, it is highly understudied. In fact, I am aware of only one study to date that has examined the question how lasting personality changes in accordance with fluctuating social environmental contexts are (Jackson et al., 2012). My third study addressed this research gap by investigating whether sojourn effects on Big Five personality trait changes might last for four years after return.

To that end, we compared sojourners and stayers in a quasi-experimental setting regarding their mean personality trait changes across a five-year time span. This allowed us to

link student sojourns as contextual predictors to relative differences in long-term personality changes at the population level. This way, I aimed at drawing conclusions on general (not person-specific) mechanisms behind individual long-term differences in personality changes (study III). Three scenarios were conceivable: First, according to the *dynamic equilibrium* (or classical set-point) *model*, effects of social environments on personality changes reflect fluctuating, and time-limited experiences (Headey & Wearing, 1989; Lykken & Tellegen, 1996), and are thus reversing. Second, (personality and) sojourns could reinforce each other (Netz & Jaksztat, 2014). The notion that trait levels and life experiences can mutually reinforce each other is known as the *corresponsive principle* (Jeronimus, Riese, Sanderman, & Ormel, 2014), possibly leading to accentuating differences between sojourners and stayers in the long run (cumulative continuity). And third, meaningful social experiences as a student sojourn could promote accelerated personality maturation, e.g., by adopting reference values (Denissen et al., 2013), with no differences between sojourners and stayers in trait changes thereafter. **Figure 4** shows these possible scenarios in long-term mean-level changes as responses to a life event as divergences from a baseline after a short-term effect has been found.

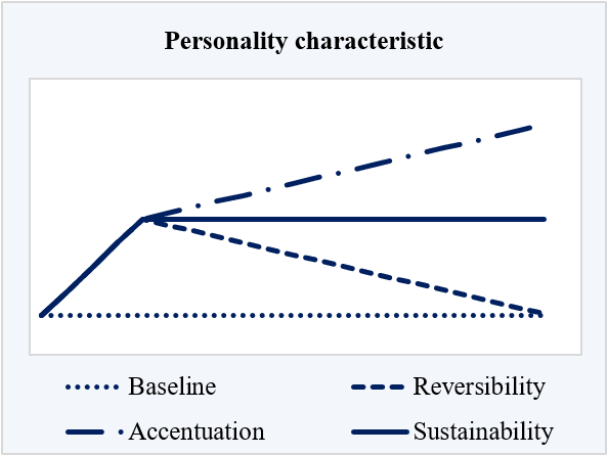

Figure 4. Possible scenarios of long-term personality change after a life event (adapted from study III).

We tested these three scenarios using data of the longitudinal project Personality Development of Sojourners (PEDES; Zimmermann & Neyer, 2013) applying multiple LCM. More specifically, we tested the predicting effect of a sojourn on trait change between T3 and T4 after controlling for its effects on the prior time intervals. Due to the reduced sample size and accordingly reduced accuracy (Kelley & Rausch, 2006) across the last time interval, we applied additional analyses to make our results more robust. Findings revealed sustained or accentuated differences in openness change, and sustained or reversed differences in agreeableness and neuroticism changes. That is, while a reversed (i.e., set-point) sojourn

effect on long-term openness change seemed unlikely according to our analyses, so did accentuated (i.e., cumulated) sojourn effects on agreeableness and neuroticism changes. This study was, to my knowledge, the first that suggested long-term effects of a left social context on openness change. Moreover, in line with findings on short-term effects of life events on personality trait changes (Bleidorn et al., 2012), this is a further hint on differences between traits in the long-term susceptibility to the environment.

According to the sociogenomic personality psychology (Roberts & Jackson, 2008), the accentuated or sustained effect on openness might be explained by higher intercultural openness in former sojourn students, which might have been triggered through more diverse intercultural (social) experiences. One explanation for such a mechanism would be a long-term shift in individuals' identity concepts (Roberts & Wood, 2006), which might be established via daily behavioral changes during the sojourn. Zimmermann and Neyer's (2013) findings on increased international relationships during a sojourn suggest that these changes might be explained by social interactions. In line with the contact hypothesis (Allport, 1954), correlational findings on diverse contact and openness to diversity support this notion (Longerbeam, 2010). These associations open the door for a highly speculative hypothesis: Might it be worthwhile to consider student sojourns as a long-term contribution to globalization? If so, this would literally put calls for personality change policies (Bleidorn et al., 2019) into social context.

Of course, we have not yet established that our findings are robust (across different samples but also within students), nor can we exclude the possibility that sojourners differ in their openness change from stayers for totally different reasons. However, these findings might be an interesting starting point for the future research and (political) debate on long-term effects of student sojourns on personality changes. From a theoretical perspective, it would be desirable to access waiting-group designs with multiple short-term measurement time points, and to include information on long-term social interactions to capture mechanisms behind our preliminary findings (see, for example, the TESSERA framework on personality changes; Wrzus, & Roberts, 2017).

IV. Conclusions From the Findings on Predictors of Personality and Personality Change

My studies were designed to answer the question which factors might in part explain individual differences in personality levels and changes. By doing so, my dissertation might help to understand personality as embedded in a social context more deeply. For example, we found strong gender differences regarding age trends in achievement and power motives. Neither age nor gender alone seemed to suffice in explaining explicit motives. If interactions of age and gender serve as proxies for social roles, these findings hint at the relevance of social roles in predicting mean levels in motives. Moreover, conflict in romantic relationships predicted subsequent changes in self-esteem, stressing the importance of social interactions for personality change.

From the perspective of a need for personality change policies (Bleidorn et al., 2019), we could deduce some hypotheses for a starting point. Self-esteem has been found to predict several mental and physical health outcomes (Rieger et al., 2016; Trzesniewski et al., 2006). Knowing that negative couple dynamics seem to enforce these outcomes by decreasing the partners' self-esteem might help to buffer or prevent these effects. Couple consultants might explicitly work on their clients' conflict behavior to prevent negative effects of conflict on one or both partners' self-esteem (study II). Besides, knowing that student sojourns might have long-term buffering effects on individuals' openness decrease (study III) might motivate politicians to foster exchange programs. Put differently, decision makers might ask whether it is possible to establish a healthy self-esteem via conflict consulting or a long lasting open-mind attitude via obligatory student sojourns. I need to stress again that it is of course necessary to first test the robustness and generalizability of our findings.

Insight into the predictors of personality change might further help people who want to change their personality traits. Many people strive for greater maturity, especially higher emotional stability (see, e.g., Robinson, Noftle, Guo, Asadi, & Zhang, 2015; Hudson & Fraley, 2015). In the light of the benefits of emotional stability (Denissen et al., 2018; Ormel et al., 2013), it is reassuring that therapy might be fruitful to promote increases in this trait (Roberts, Luo, Briley, Chow, Su, & Hill, 2017). Thus, it seems possible to intentionally change in personality. The knowledge on possible long-term effects of student sojourns further implies that a temporary change in the social context might be able to foster long-term personality change.

Moreover, people differ in their capacities to effectively change their personality (Hennecke, Bleidorn, Denissen, & Wood, 2014), which are themselves associated with their personality. While this might be sobering, "the more, the better" is not always true. For example, Denissen et al. (2018) found that higher levels of conscientiousness and agreeableness can be detrimental to income when individuals are placed in a workplace where these traits do not need to be high. Niehoff et al. (2017) found that individuals with the highest conscientiousness and agreeableness levels changed these traits to be more average when going abroad (in contrast to the generally positive sojourn effects on these traits). Given these were not artificial findings due to regression to the mean, these individuals' personality change might also indicate an adaptation to their context. Although we can of course not conclude this at this point, (some) people might be able to purposefully (not necessarily intentionally) change in the direction of more useful behaviors. From this perspective, it might be worthwhile to not only address personality change from the standpoint of an optimal personality level (e.g., as conscientious as possible), but to try to understand personality as related to the social context in which it unfolds.

One hint at this notion has been derived from study II. Supporting the sociometer perspective, our findings implied that individuals' self-esteem levels serve as indicators to assess how well they are doing socially. With this function, the optimal level of self-esteem is defined by an individual's need to improve their social functioning. Following, the (optimal) self-esteem level per person is directly dependent on the individual's social context. Study I provided further support for the idea that there is no general ideal level of any personality characteristic. Instead, we found that age and gender interact to predict mean levels in explicit motives. Combining theories on lifespan adaptation processes with social role demands, it seems logical to conclude that levels of explicit motives are indicative of an adaptation to the social context. Following, it could be true that individuals are able to adapt their personality more purposefully than we have previously thought. Please note that this does not necessarily mean that this process is conscious. However, in line with Denissen et al. (2018), future studies might invest more time in testing whether societally desired levels of personality characteristics are indeed a sign for adaptation across different contexts, or if this might in some cases be a misconception.

While my studies have provided new insights in personality change across social contexts, of course, there are still open questions. One important question is whether the found effects can indeed be attributed to environmental influences. More recent extended

behavioral genetic studies make it possible to model social contextual variables (e.g., relationship conflict) as moderators of differences in personality levels (Krueger, South, Johnson, & Iacono, 2009). This way, they not only enable us to consider the relative predictive effects of genes and the environment per se, but also their relative impact at varying levels of the moderator. Using this approach, Johnson and Krueger (2006) as well as Krueger et al. (2009) found that the heritability of a personality characteristic can vary depending on the level of a social context variable. For example, the heritability of 17-year-old adolescents' negative emotionality decreased as a function of their perceived conflict frequency with their parents – that is, the more frequently the conflict, the lower its estimated heritability, and vice versa. While this kind of approach allows to understand cross-sectional gene–environment interactions more deeply, it still does not yield information on developmental processes, for example on the question *which* individuals change in accordance with these factors and why.

There is an increasing body of research implying that age or a particular life event does not have the same (amount of) impact on everybody (e.g., Denissen et al., 2019; Graham et al., 2020; Specht et al., 2011). For example, while retirement has been found to be associated with mean-level changes in some Big Five traits, there were significant individual differences in the response to this life event (Schwaba & Bleidorn, 2019). We found the same pattern in study II and III. Bleidorn and colleagues (2020) currently argue for longitudinal studies that incorporate daily assessments on the basis of everyday-life experiences. With these methods, it might be possible to measure interactions of personality characteristics with each other and with contextual factors, eventually disentangling processes of personality changes. Besides the notion that a characteristic's heritability can vary depending on the social context, susceptibility to the environment in which an individual is situated is at least in part heritable (Del Giudice, 2016; Richards et al., 2017). To capture both genetic and environmental effects in a dynamic way, it is moreover necessary to model heritability and the influence of specific environmental factors longitudinally in the same analysis. An exciting approach is the combination of these techniques. If this succeeds, we might eventually be able to explain reasons and benefits of particular personality levels and changes in varying social contexts.

References

Allport, G. W. (1954). *The nature of prejudice*. Cambridge, Mass: Addison-Wesley Pub. Co.

Bleidorn, W., Hill, P., Back, M. D., Denissen, J. J. A., Hennecke, M., Hopwood, C. J., … Roberts, B. W. (2019). The policy relevance of personality traits. *American Psychologist, 74,* 1056–1067. https://doi.org/10.1037/amp0000503

Bleidorn, W., Hopwood, C., Back, M. D., Denissen, J. J. A., Hennecke, M., Jokela, M., … Zimmermann, J. (2020). Longitudinal Experience-Wide Association Studies (LEWAS) – A framework for studying personality change. *European Journal of Personality*. https://doi.org/10.31234/osf.io/krfw9

Bleidorn, W., Kandler, C., Hülsheger, U. R., Riemann, R., Angleitner, A., & Spinath, F. M. (2010). Nature and nurture of the interplay between personality traits and major life goals. *Journal of Personality and Social Psychology, 99,* 366–379. https://doi.org/10.1037/a0019982

Bleidorn, W., Kandler, C., Riemann, R., Spinath, F. M., & Angleitner, A. (2009). Patterns and sources of adult personality development: Growth curve analyses of the NEO PI-R scales in a longitudinal twin study. *Journal of Personality and Social Psychology, 97,* 142–155. https://doi.org/10.1037/a0015434

Brüderl, J., Drobnič, S., Hank, K., Huinink, J., Nauck, B., Neyer, F. J., … Wilhelm, B. (2018). The German Family Panel (pairfam). GESIS Data Archive, Cologne. ZA5678 Data file Version 9.1.0, https://doi.org/10.4232/pairfam.5678.9.1.0

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

Carstensen, L. L. (2006). The influence of a sense of time on human development. *Science, 312,* 1913–1915. https://doi.org/10.1126/science.1127488

Caspi, A., & Roberts, B. W. (2001). Personality development across the life course: The argument for change and continuity. *Psychological Inquiry*, 12, 49–66. https://doi.org/10.1207/S15327965PLI1202_01

Costa, P. T., Jr., Terracciano, A., & McCrae, R. R. (2001). Gender differences in personality traits across cultures: Robust and surprising findings. *Journal of Personality and Social Psychology, 81*, 322–331. https://doi.org/10.1037/0022-3514.81.2.322

DAAD (2015). DAAD-IW-Studie 2015 [DAAD-IW study 2015]. Retrieved from https://eu.daad.de/service/auswertung-und-statistik/studien-und-auswertungen-der-na-daad/daad-iw-studie-2015/de/49310-daad-iw-studie-01/

Del Giudice, M. (2016). Differential susceptibility to the environment: Are developmental models compatible with the evidence from twin studies? *Developmental Psychology, 52,* 1330–1339. https://doi.org/10.1037/dev0000153

Denissen, J. J. A., Bleidorn, W., Hennecke, M., Luhmann, M., Orth, U., Specht, J., & Zimmermann, J. (2018). Uncovering the power of personality to shape income. *Psychological Science, 29*, 3–13. https://doi.org/10.1177/0956797617724435

Denissen, J. J. A., Luhmann, M., Chung, J. M., & Bleidorn, W. (2019). Transactions between life events and personality traits across the adult lifespan. *Journal of Personality and Social Psychology, 116*, 612–633. https://doi.org/10.1037/pspp0000196

Denissen, J. J. A., Ulferts, H., Lüdtke, O., Muck, P. M., & Gerstorf, D. (2014). Longitudinal transactions between personality and occupational roles: A large and heterogeneous study of job beginners, stayers, and changers. *Developmental Psychology, 50*, 1931–1942. https://doi.org/10.1037/a0036994

Denissen, J. J. A., van Aken, M. A. G., Penke, L., & Wood, D. (2013). Self-regulation underlies temperament and personality: An integrative developmental framework. *Child Development Perspectives, 7*, 255–260. https://doi.org/10.1111/cdep.12050

Denissen, J. J. A., van Aken, M. A. G., & Roberts, B. W. (2011). Personality development across the life span. In T. Chamorro-Premuzic, S. von Stumm, & A. Furnham (Eds.), *The Wiley-Blackwell handbooks of personality and individual differences* (p. 77–100). Chichester, UK: Wiley-Blackwell.

Dobewall, H., & Aavik, T. (2016). Rank-order consistency and profile stability of self- and informant-reports of personal values in comparison to personality traits. *Journal of Individual Differences, 37*, 40–48. https://doi.org/10.1027/1614-0001/a000186

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

Enders, C. K. (2011). Analyzing longitudinal data with missing values. *Rehabilitation Psychology, 56*, 267–288. https://doi.org/10.1037/a0025579

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

Fletcher, G. J. O., Simpson, J. A., & Thomas, G. (2000). The measurement of perceived relationship quality components: A confirmatory factor analytic approach. *Personality and Social Psychology Bulletin, 26*, 340–354. https://doi.org/10.1177/0146167200265007

Funder, D. C., & Sneed, C. D. (1993). Behavioral manifestations of personality: An ecological approach to judgmental accuracy. *Journal of Personality and Social Psychology, 64*, 479–490. https://doi.org/10.1037/0022-3514.64.3.479

Graham, E. K., Weston, S. J., Gerstorf, D., Yoneda, T. B., Booth, T., Beam, C. R., . . . Mroczek, D. K. (2020). Trajectories of big five personality traits: A coordinated analysis of 16 longitudinal samples. *European Journal of Personality*, https://doi.org/10.1002/per.2259

Greischel, H., Noack, P., & Neyer, F. J. (2016). Sailing uncharted waters: Adolescent personality development and social relationship experiences during a year abroad. *Journal of Youth and Adolescence*, *45*, 2307–2320. https://doi.org/10.1007/s10964-016-0479-1

Harris, M. A., & Orth, U. (2019). The link between self-esteem and social relationships: A meta-analysis of longitudinal studies. *Journal of Personality and Social Psychology*. https://doi.org/10.1037/pspp0000265

Headey, B., & Wearing, A. (1989). Personality, life events, and subjective well-being: Toward a dynamic equilibrium model. *Journal of Personality and Social Psychology*, *57*, 731–739. https://doi.org/10.1037/0022-3514.57.4.731

Heckhausen, J. (1997). Developmental regulation across adulthood: Primary and secondary control of age-related challenges. *Developmental Psychology*, *33*, 176–187. https://doi.org/10.1037/0012-1649.33.1.176

Hennecke, M., Bleidorn, W., Denissen, J. J. A., & Wood, D. (2014). A three-part framework for self-regulated personality development across adulthood. *European Journal of Personality*, *2*8, 289–299. https://doi.org/10.1002/per.1945

Helson, R., Kwan, S., John, O. P., & Jones, C. (2002). The growing evidence for personality change in adulthood: Findings from research with personality inventories. *Journal of Research in Personality*, *36*, 287–306. https://doi.org/10.1016/S0092-6566(02)00010-7

Hudson, N. W., & Fraley, R. C. (2015). Volitional personality trait change: Can people choose to change their personality traits? *Journal of Personality and Social Psychology*, *109*, 490–507. https://doi.org/10.1037/pspp0000021

Huinink, J., Brüderl, J., Nauck, B., Walper, S., Castiglioni, L., & Feldhaus, M. (2011). Panel Analysis of Intimate Relationships and Family Dynamics (pairfam): Conceptual framework and design. *Journal of Family Research, 23*, 77–101.

Jackson, J. J., Thoemmes, F., Jonkmann, K., Lüdtke, O., & Trautwein, U. (2012). Military training and personality trait development: Does the military make the man, or does the man make the military? *Psychological Science*, *23*, 270–277. https://doi.org/10.1177/0956797611423545

Jeronimus, B. F., Riese, H., Sanderman, R., & Ormel, J. (2014). Mutual reinforcement between neuroticism and life experiences: A five-wave, 16-year study to test reciprocal causation. *Journal of Personality and Social Psychology, 107*, 751–764. https://doi.org/10.1037/a0037009

Johnson, W., & Krueger, R. F. (2006). How money buys happiness: Genetic and environmental processes linking finances and life satisfaction. *Journal of Personality and Social Psychology, 90*, 680–691. https://doi.org/10.1037/0022-3514.90.4.680

Kandler, C., Penner, A., Richter, J., & Zapko-Willmes, A. (2019). The Study of Personality Architecture and Dynamics (SPeADy): A longitudinal and extended twin family study. *Twin Research and Human Genetics, 1*, 1–6. https://doi.org/10.1017/thg.2019.62

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,* 231–243. https://doi.org/10.1002/per.1952

Kelley, K., & Rausch, J. R. (2006). Sample size planning for the standardized mean difference: Accuracy in parameter estimation via narrow confidence intervals. *Psychological Methods, 11*, 363–385. https://doi.org/10.1037/1082-989X.11.4.363

Kenny, D. A., Kashy, D. A., & Cook, W. L. (2006). Dyadic data analysis. New York, NY: Guilford Press. Retrieved from https://books.google.de/books?hl=de&lr=&id=D7OyERQVQ1cC&oi=fnd&pg=PA1&dq=dyadic+data+analysis&ots=UCxZmI4iHt&sig=9gKlaTwAmA0Dq6_aC4PJivDBcUo#v=onepage&q=dyadic%20data%20analysis&f=false

Kim, H., Di Domenico, S. I., & Connelly, B. S. (2018). Self–other agreement in personality reports: A meta-analytic comparison of self- and informant-report means. *Psychological Science*, *30*, 1–10, https://doi.org/10.1177/0956797618810000

Krueger, R. F., South, S., Johnson, W., & Iacono, W. (2009). The heritability of personality is not always 50%: Gene–environment interactions and correlations between personality and parenting. Journal of Personality, 76, 1485–1521. https://doi.org/10.1111/j.1467-6494.2008.00529.x

Leary, M. R., Tambor, E. S., Terdal, S. K., & Downs, D. L. (1995). Self-esteem as an interpersonal monitor: The sociometer hypothesis. *Journal of Personality and Social Psychology, 68*, 518–530. https://doi.org/10.1037/0022-3514.68.3.518

Lehnart, J., Neyer, F. J., & Eccles, J. (2010). Long-term effects of social investment: The case of partnering in young adulthood. *Journal of Personality, 78*, 639–670. https://doi.org/10.1111/j.1467-6494.2010.00629.x

Lewis, M. (2001). Issues in the study of personality development. *Psychological Inquiry, 12*, 67–83. https://doi.org/10.1207/S15327965PLI1202_02

Longerbeam, S. (2010). Developing openness to diversity in living-learning program participants. *Journal of Diversity in Higher Education, 3*, 201–217. https://doi.org/10.1037/a0021277

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, 620–637.* https://doi.org/10.1037/a0023743

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

Magnusson, D. (1990). Personality development from an interactional perspective. In L. Pervin (Ed.), *Handbook of personality: Theory and research* (pp. 193–222). New York, NY: New York: Guilford Press.

McAdams, D. P. (2013). The psychological self as actor, agent, and author. *Perspectives on Psychological Science, 8*, 272–295. https://doi.org/0.1177/1745691612464657

McAdams, D. P., & Olson, B. D. (2010). Personality development: Continuity and change over the life course. *Annual Review of Psychology, 61*, 517–42. https://doi.org/10.1146/annurev.psych.093008.100507

McArdle, J. J. (2009). Latent variable modeling of differences and changes with longitudinal data. *Annual Review of Psychology, 60*, 577–605. https://doi.org/10.1146/annurev.psych.60.110707.163612

McCrae, R. R., & Costa, P. T., Jr. (1999). *A Five-Factor theory of personality.* In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (p. 139–153). Guilford Press.

McCrae, R. R., Costa, P. T., Jr., de Lima, M. P., Simões, A., Ostendorf, F., Angleitner, A., … Piedmont, R. L. (1999). Age differences in personality across the adult life span: Parallels in five cultures. *Developmental Psychology, 35*, 466–477. https://doi.org/10.1037/0012-1649.35.2.466

McCrae, R. R., Costa, P. T., Jr., Ostendorf, F., Angleitner, A., Hřebíčková, M., Avia, M. D.,…Smith, P. B. (2000). Nature over nurture: Temperament, personality, and life span development. *Journal of Personality and Social Psychology, 78*, 173–186. https://doi.org/10.1037/0022-3514.78.1.173

Merkel, A. D. (2020, March 07). *Potenzial von Männern und Frauen gleichermaßen nutzen* [Video podcast].

Mund, M., Finn, C., Hagemeyer, B., & Neyer, F. J. (2016). Understanding dynamic transactions between personality traits and partner relationships. *Current Directions in Psychological Science, 25*, 411–416. https://doi.org/10.1177/0963721416659458

Mund, M., Finn, C., Hagemeyer, B., Zimmermann, J., & Neyer, F. J. (2015). The dynamics of self-esteem in partner relationships. *European Journal of Personality, 28*, 235–249. https://doi.org/10.1002/per.1984

Netz, N., & Jaksztat, S. (2014). Mobilised by mobility? Determinants of international mobility plans among doctoral candidates in Germany. In M. Tight & N. Maadad (Eds.), *International perspectives on higher education research: Academic mobility* (Vol. 11, pp. 35–59). Bingley, U.K: Emerald. https://doi.org/10.1108/S1479-362820140000011009

Neyer, F. J., & Asendorpf, J. (2018). *Psychologie der Persönlichkeit* (6th edition). Berlin, Germany: Springer.

Neyer, F. J., & Lehnart, J. (2007). Relationships matter in personality development: Evidence from an 8-year longitudinal study across young adulthood. *Journal of Personality, 75*, 535–568. https://doi.org/10.1111/j.1467-6494.2007.00448.x

Niehoff, E., Petersdotter, L., & Freund, P. A. (2017). International sojourn experience and personality development: Selection and socialization effects of studying abroad and the Big Five. *Personality and Individual Differences, 112*, 55–61. https://doi.org/10.1016/j.paid.2017.02.043

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

Peterson Tilden, V. & Galyen, R. D. (1987). Cost and conflict: The darker side of social support. *Western Journal of Nursing Research, 9*, 9–18. https://doi.org/10.1177/019394598700900103

Richards, J. S., Vásquez, A. A., van Rooij, D., van der Meer, D., Franke, B., Hoekstra, P. J., … Buitelaar, J. K. (2017). Testing differential susceptibility: Plasticity genes, the social environment, and their interplay in adolescent response inhibition. *The World Journal of Biological Psychiatry*, *18*, 308–321. https://doi.org/10.3109/15622975.2016.1173724

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*, e16–e22. https://doi.org/10.1037/pspp0000037

Roberts, B. W., & Jackson, J. J. (2008). Sociogenomic personality psychology. *Journal of Personality, 76*, 1523–1544. https://doi.org/10.1111/j.1467-6494.2008.00530.x

Roberts, B. W., Luo, J., Briley, D. A., Chow, P. I., Su, R., & Hill, P. L. (2017). A systematic review of personality trait change through intervention. *Psychological Bulletin, 143*, 117–141. https://doi.org/10.1037/bul0000088

Roberts, B. W., & Mroczek, D. K. (2008). Personality trait change in adulthood. *Current Directions in Psychological Science, 17*, 31–35. https://doi.org/10.1111/j.1467-8721.2008.00543.x

Roberts, B. W., & Robins, R. W. (2000). Broad dispositions, broad aspirations: The intersection of personality traits and major life goals. *Personality and Social Psychology Bulletin, 26*, 1284–1296. https://doi.org/10.1177/0146167200262009

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

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

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* (3rd edition, pp. 375–398). New York, NY: Guilford.

Roberts, B. W., Wood, D., & Smith, J. L. (2005). Evaluating Five Factor theory and social investment perspectives on personality trait development. *Journal of Research in Personality, 39*, 166–184. https://doi.org/10.1016/j.jrp.2004.08.002

Robinson, O. C., Noftle, E. E., Guo, J., Asadi, S., Zhang, X. (2015). Goals and plans for Big Five personality trait change in young adults. *Journal of Research in Personality, 59*, 31–43, https://doi.org/10.1016/j.jrp.2015.08.002

Schmitt, D. P., Realo, A., Voracek, M., & Allik, J. (2008). Why can't a man be more like a woman? Sex differences in Big Five personality traits across 55 cultures. *Journal of Personality and Social Psychology, 94*, 168–182. https://doi.org/10.1037/0022-3514.94.1.168

Schwaba, T., & Bleidorn, W. (2019). Personality trait development across the transition to retirement. *Journal of Personality and Social Psychology, 116*, 651–665. https://doi.org/10.1037/pspp0000179

Specht, J., Egloff, B., & Schmukle, S. C. (2011). Stability and change of personality across the life course: The impact of age and major life events on mean-level and rank-order stability of the Big Five. *Journal of Personality and Social Psychology, 101*, 862–882. https://doi.org/10.1037/a0024950

Specht, J., Bleidorn, W., Denissen, J. J. A., Hennecke, M., Hutteman, R., Kandler, C., … Zimmermann, J. (2014). What drives adult personality development? A comparison of theoretical perspectives and empirical evidence. *European Journal of Personality, 28*, 216–230. https://doi.org/10.1002/per.1966

Srivastava, S., & Beer, J. S. (2005). How self-evaluations relate to being liked by others: Integrating sociometer and attachment perspectives. *Journal of Personality and Social Psychology, 89*, 966–977. https://doi.org/10.1037/0022-3514.89.6.966

Steyer, R., Eid, M. & Schwenkmezger, P. (1997). Modeling true intraindividual change: True change as a latent variable. *Methods of Psychological Research Online, 2*, 21–33. Retrieved from https://www.dgps.de/fachgruppen/methoden/mpr-online/issue2/art2/steyer.pdf

Steyer, R., Mayer, A., Geiser, C., & Cole, D. A. (2015). A theory of states and traits—Revised. *Annual Review of Clinical Psychology, 11*, 71–98. https://doi.org/10.1146/annurev-clinpsy-032813-153719

Steyer, R., Schmitt, M. & Eid, M. (2000). Latent state-trait theory and research in personality and individual differences. *European Journal of Personality, 13*, 389-408. https://doi.org/10.1002/(SICI)1099-0984(199909/10)13:5<389::AID-PER361>3.0.CO;2-A

Trzesniewski, K. H., Donnellan, M. B., Moffitt, T. E., Robins, R. W., Poulton, R., Caspi, A (2006). Low self-esteem during adolescence predicts poor health, criminal behavior, and limited

economic prospects during adulthood. *Developmental Psychology, 42*, 381–90. https://doi.org/10.1037/0012-1649.42.2.381

Wagerman, S. A., & Funder, D. C. (2007). Acquaintance reports of personality and academic achievement: A case for conscientiousness. *Journal of Research in Personality, 41*, 221–229. https://doi.org/10.1016/j.jrp.2006.03.001

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*, 148–163. https://doi.org/10.1037/a0029680

Watson, J. B. (1930). *Behaviorism*. Chicago: University of Chicago Press.

Wrzus, C., & Roberts, B. W. (2017). Processes of personality development in adulthood: The TESSERA framework. *Personality and Social Psychology Review: An Official Journal of the Society for Personality and Social Psychology, Inc, 21*, 253–277. https://doi.org/10.1177/1088868316652279

Zimmermann, J., & Neyer, F. J. (2013). Do we become a different person when hitting the road? Personality development of sojourners. *Journal of Personality and Social Psychology, 105*, 515–530. https://doi.org/10.1037/a0033019

Author Contributions

Below, I am listing the respective contributions of each author that was involved in the process of planning, exerting, and/or writing of one or more of the manuscripts presented in this dissertation. All involved authors approved of these classifications.

I. Age trends in explicit motives across the lifespan — Insights from self- and informant reports.

Conceptualization:	Julia Richter
Data Collection:	Julia Richter, Alexandra Zapko-Willmes
Formal Analysis:	Julia Richter
Methodology:	Julia Richter
Resources:	Christian Kandler
Supervision:	Christian Kandler
Visualization:	Julia Richter
Writing – Original Draft:	Julia Richter
Writing – Review & Editing:	Julia Richter, Alexandra Zapko-Willmes, Christian Kandler

II. Transactions between self-esteem and conflict in romantic relationships — A five-year longitudinal study.

Conceptualization:	Julia Richter
Data Curation:	Julia Richter, Christine Finn
Formal Analysis:	Julia Richter
Methodology:	Julia Richter, Christine Finn
Resources:	Christine Finn
Supervision:	Julia Richter, Christine Finn
Visualization:	Julia Richter, Christine Finn
Writing – Original Draft:	Julia Richter
Writing – Review & Editing:	Julia Richter, Christine Finn

III. Do sojourn effects on personality trait changes last? A five-year longitudinal study.

Conceptualization:	Julia Richter
Data Collection & Curation:	Julia Richter, Julia Zimmermann
Formal Analysis:	Julia Richter
Methodology:	Julia Richter, Christian Kandler, Julia Zimmermann
Resources:	Christian Kandler, Julia Zimmermann, Franz J. Neyer
Supervision:	Christian Kandler, Julia Zimmermann, Franz J. Neyer
Visualization:	Julia Richter
Writing – Original Draft:	Julia Richter
Writing – Review & Editing:	Julia Richter, Christian Kandler, Julia Zimmermann, Franz J. Neyer

Comments on the Appendices

Below, you will find the manuscripts submitted or accepted for publication in the order that I referred to them during the synopsis. To facilitate readability, I have included Tables and Figures directly in the body of the manuscripts close to the place where they are first mentioned. Moreover, for two manuscripts I adapted the headlines to make them more comparable. I also altered the formats of Figures and Tables to make them fit more easily on the pages. These formatting adaptions have not altered the contents of my manuscripts in any form.

Supplemental materials can either be found at the Open Science Framework via the provided links that are directly implemented in the manuscripts or below the reference list, depending on the procedure implemented for the respective journal.

Appendix I

Age Trends in Explicit Motives Across the Lifespan — Insights From Self- and Informant Reports

Julia Richter, Alexandra Zapko-Willmes, & Christian Kandler

Abstract

Although explicit motives are said to underlie goals across the lifespan, we know little about their age trends. This study tested lifespan theories' proposition that motives shift from personal success and growth towards fewer but closer relationships and protection. We expected higher levels of affiliation, achievement, power, and personal growth, and lower levels of intimacy and health protection among younger compared to older adults. Self-reports (cross-sectional German-speaking sample; N = 2,278; 14–94 years; 61.6% females) and ratings from well-informed acquaintances (n = 944; 41.4%) and their combination in a multivariate latent structure model allowed an analysis of females' and males' linear and nonlinear age differences. Most hypotheses were met, especially for elderly people. Males' achievement and power showed no age associations.

Key words: explicit motives; self-reports; informant reports; multi-rater study; age differences; gender differences

Theory and Review on Age Trends in Explicit Motives Across the Lifespan

Introduction

Aging people perceive increasingly less control over their goals by noticing physical and cognitive decline as indicators of their limited remaining life time (Heckhausen, Dixon, & Baltes, 1989). Although perceived loss of control over committing to and pursuing certain goals is said to jeopardize a positive view of the self (Heckhausen, 1997), older adults seem to be as equally adapted as younger adults in terms of self-concept and mood (e.g., Brandtstädter, Wentura, & Greve, 1993; Mroczek, 2001; Rothermund, & Brandtstädter, 2003). One reason for this might be that human beings change their motivational focus and strategies with age to optimize well-being, shifting away from motives and goals related to personal success and expansion towards the focus on maintaining functionality and emotional well-being (e.g., Brandtstädter, Rothermund, Kranz, & Kühn, 2010; Carstensen, 2006; Freund, Hennecke, & Riedinger, 2008; Heckhausen, 1997).

Previous findings on life goals have mainly provided support for this hypothesis (e.g., Bühler, Weidmann, Nikitin, & Grob, 2019; Ebner, Freund, & Baltes, 2006). These findings were, however, only based on self-ratings, and did not take gender differences into consideration. The aim of our study thus was to test if predictions of lifespan theories regarding age trends in six explicit motives (i.e., motives that are accessible to the conscious mind) hold not only for self-raters, but also for reports from well-informed acquaintances and across genders. Moreover, we combined rater perspectives to yield more accurate information on motives, controlled for variance due to self-rater biases (e.g., self-deception, impression management) as well as other rater-specific biases (e.g., leniency or severity effects). We assumed that expansion-oriented motives (i.e., affiliation, achievement, power, and personal growth) are less important for older compared to younger people, whereas protection-oriented motives (i.e., intimacy and health protection) are more important. We used a large age-heterogeneous multi-rater sample with an age range of about 80 years to test this expansion-to-protection shift hypothesis.

Classifying Motives Across the Lifespan as Expansive vs. Protective

Literature has often mixed up various motivation-related psychological constructs, such as motives, major life goals, personal projects, life tasks, and value priorities (e.g., Davidov, Schmidt, & Schwartz, 2008; McAdams, 2013). Schönbrodt and Gerstenberg (2012) used an item-response theory approach to integrate these constructs within a comprehensive

framework. They subsumed "personal goals, abstract values, personality attributes, interests, attitudes, and affective preferences" (p. 732) under the phrase "*explicit motives*". Explicit motives reflect conscious underlying tendencies towards multiple motive-related outcomes, such as achievement, status, intimacy, or social integration (Bleidorn, Kandler, Hülsheger, Riemann, Angleitner, & Spinath, 2010). In contrast, very specific personal or midlevel goals (e.g., graduation, having children, or living in a lakeside villa, compare Ebner et al., 2006) reflect concrete outcomes that people strive for during their lives (Roberts & Robins, 2000). Explicit motives as "self-attributed" and "conscious" (McClelland, Koestner, & Weinberger, 1989, p. 690) differ from implicit motives (Hagemeyer, Dufner, & Denissen, 2016; McClelland et al., 1989). This has been supported by generally low correlations across the two (Köllner & Schultheiss, 2014; Schultheiss & Brunstein, 2001). Still, terms for implicit and explicit motives are sometimes identical (compare Denzinger, Backes, Job, & Brandstätter, 2016, with Schönbrodt & Gerstenberg, 2012).

With regard to resource orientation, it is useful to organize motives as *expansive* or *protective* (compare Carstensen, 2006; Lang, 2000; Ebner et al., 2006; see also Freund, 2006). Achievement and power motives are defined as the will to expand knowledge and influence (Schönbrodt & Gerstenberg, 2012). According to the socioemotional selectivity theory (Carstensen, 2006; Carstensen, Isaacowitz, & Charles, 1999), affiliation is also geared towards expanding (social) resources. Specifically, raising the quantity of acquaintances or having contact to a lot of various people can serve the goal to collect information. Intimacy, by contrast, is related to maintaining and protecting existing close relationships, which are defined by their high emotional support, providing well-being due to long-established connections (Carstensen, 2006; Kasser & Ryan, 1996; Lang, 2000). In line with this differentiation, social network size seems to decrease across age, while the rated quality of close relationships does not increase or even increases (Lang, 2000; Wrzus, Hänel, Wagner, & Neyer, 2013). From both a core human motivational and a lifespan perspective, two further explicit motives are worthwhile to consider: personal growth and health protection. Personal growth enhances personal maturity (Schwartz et al., 2012), and, thus, expands resources, whereas health protection reflects an orientation towards the maintenance (in contrast to the improvement) of one's own health and physical fitness (Klusmann, Trautwein, & Lüdtke, 2005), and is thus inherently protective.

Theories on Age-Related Shifts in Explicit Motives

Lifespan researchers have argued for a shift in focus from expansion to protection and loss avoidance with age (Ebner et al., 2006; Lang, 2000; McAdams, 2013). Put differently, the importance of protective motives has been assumed to increase, and that of expansive motives to decrease (Brandtstädter et al., 2010; Ritter & Freund, 2014). The hypothesized age-related trends in motives may be linked with the necessity to change control strategies towards a more pronounced focus on avoidance of losses in comparison to developmental gains (e.g., Heckhausen, 1989, 1997; Heckhausen & Schulz, 1995). For example, individuals might face poorer health and physical functionality at an older (compared to younger) age. Thus, they might now stay home with their family and friends and avoid risky behavior to protect their health, when in former days, they had loved to do adventurous things and meet new people.

According to the socioemotional selectivity theory, age trends in motives are mainly due to the subjective perception of time (Carstensen et al., 1999). Perceiving remaining life time to be limited leads to an orientation towards "finding emotional meaning and satisfaction from life and invest fewer resources into gathering information and expanding horizons" (Carstensen, 2006, p. 1915; see also Freund et al., 2008). That is, older adults (in comparison to younger adults) focus on protecting close relationships that provide emotional support instead of widening their network (Lang, 2000). In contrast, perceiving time as seemingly unlimited, as healthy adolescents and young adults often do, may result in the less timely rewarding pursuit of expansive goals (Brandtstädter et al., 2010; Carstensen & Fredrickson, 1998), like financial success (Borg, Hertel, & Hermann, 2017) or expanding the social network (Carstensen, 2006). Older people may also have to invest more to maintain their established resources than younger people (Freund & Ebner, 2005; Heckhausen, Wrosch, & Schulz, 2010). Considering the need to balance resource demands (Ebner et al., 2006), it may be reasonable that the need to stabilize these resources becomes more important than to further expand them. The latter also makes sense in view of limited expected remaining lifetime to use available resources (Baltes & Baltes, 1990).

Empirical Findings on Age Trends in Explicit Motives

Studies on age trends in a broad range of explicit motives are rare. Bühler et al. (2019) recently examined age trends in motives as captured by the German version of the Aspirations Index (Klusmann et al., 2005). The authors found that age was negatively associated with importance and attainability of personal growth, status, and work goals, but positively with importance and attainability of prosocial-engagement goals. Moreover, the authors reported a

small positive linear age trend in health protection. These findings match lifespan theories that propose a trend from expansion towards protection motivation in adulthood (Carstensen, 2006; Ebner et al., 2006).

Studies on specific self-generated life goals are more frequent than those on abstract motives. Ebner and colleagues (2006), for example, have asked participants to write down their six most important specific goals. In a second step, participants were asked to independently rate each goal regarding each of the three orientations growth, maintenance, and loss prevention on an 8-point scale. The authors found an orientation towards growth in younger adults, and towards maintenance and loss prevention in older adults. Consistently, Nurmi (1992) found that age predicted a shift in goal importance. Younger adults mentioned self-related and future-oriented goals, whereas elderly people mentioned their own health and present-oriented goals more frequently. To our knowledge, Dunlop, Bannon, and McAdams (2017) were the first who studied mean-level and rank-order consistencies of broad goal categories in young (on average 20 years old) and midlife (between 55 and 57 years old) adults. Participants were asked to write down seven goals that were important to them, which were later coded into one of the eight conceptual categories achievement, affiliation, intimacy, power, generativity, health, finance, and travel. The mean frequency of young adults' goal categories did not significantly change across the three-year time span. In midlife adults, in contrast, the mean frequency of the motive-based goal category affiliation increased and power tended to decrease. Although the samples have not been compared directly, descriptive statistics suggested a higher mean frequency of the motive-based goal categories achievement, affiliation, and power in the younger sample, and a higher mean frequency of intimacy and health protection in the older sample.

Interestingly, no study to date has captured gender differences in age associations of motives. This is especially surprising as research on values has shown that females are more oriented towards self-transcendence (e.g., valuing welfare of significant others), while males show higher mean levels of achievement and power values (Robinson, 2013). Although explicit motives are conceptually distinct from value priorities regarding measurement method and content (e.g., Hofer, Busch, Bond, Li, & Law, 2010; McAdams, 2013), they show substantial empirical convergence at a very general level (Schönbrodt & Gerstenberg, 2012), making it probable that gender differences would also occur in motives.

To sum up, from a lifespan perspective, research has established a tendency of higher levels in achievement, affiliation, personal growth, and power goals in younger as compared

to older cohorts; while the opposite tends to be true for health protection and intimacy goals. Little is known about the role gender plays in shaping motives across the lifespan.

The Current Study

The Value of Multiple Rater Perspectives

Most previous studies were solely based on self-reports. A multi-methodological approach, however, may prove to be vital in examining explicit motives. Well-informed acquaintances, for instance, provide additional reliable (Dobewall & Aavik, 2016) and (incrementally) valid (Kim, Di Domenico, & Connelly, 2018; Wagerman & Funder, 2007) information on individual characteristics. Whereas uninformed observers may have limited access to the target person's introspection, well-acquainted informants (e.g., partners, friends, and family members) have access to multiple sources of information on the target. For example, the target's engagement, investment, decisions, and statements may all serve as important cues for action motivation. This implies that the target and their well-acquainted informants can base their evaluation on the target's motives on a broad range of the same cues. Importantly, however, informants should not feel the need to apply impression management techniques (e.g., regarding the motive to exert influence). Moreover, multi-rater data enable the sensitivity to detect perspective-specific variance in motives that is observable by close acquaintances, but not necessarily conscious to the self-rater due to, for example, self-deception (Wagerman & Funder, 2007).

Of course, informant reports may also be biased by the informants' subjective interpretations and expectations. The combined analysis of multiple informant reports in addition to self-reports is thus important to control for systematic measurement artifacts due to rater biases (e.g., halo, leniency, or severity effects). In total, the combination of self- and informant reports allows more elaborate assessments by controlling systematic response biases in both self- and informant reports at the same time. Put differently, the focus on self-other agreement ensures more accurate measures in terms of reliability and convergent validity because latent variables based on self-other covariance can neither be due to random measurement error nor due to nonrandom biases of a specific rater perspective (Funder & Sneed, 1993).

The Lifespan Perspective

As far as we know, no study to date has captured age trends in a broad range of explicit motives across the lifespan separately for females and males, and from the

perspectives of both the self and well-informed close acquaintances. With the present study, we wanted to address this gap with data from participants between 14 to 94 years of age.[2] The use of both self- and informant reports allowed us to examine age trends separately for each rater-perspective, and to apply a latent modeling approach with more accurate reflections of motives based on self-other covariance. Moreover, examining age differences separately for females and males made it possible to explore gender differences in age-related trends.

Hypotheses

Hypothesis 1: Intimacy and affiliation show higher levels in females, and achievement and power show higher levels in males.

We argue that gender plays a role in shaping age associations with motives across the lifespan (compare Robinson, 2013). We assessed all age trends separately for females and males to examine age × gender interactions.

Hypothesis 2: Health protection and intimacy are positively associated with age.

We argue that prioritizing one's own health and being close to one's most important and beloved people can be seen as expressions of protecting one's own physical and social status (e.g., Fredrickson & Carstensen, 1990; Lang, 2000; Wrzus et al., 2013). We thus expected health protection and intimacy as protective motives to be positively associated with age.

Hypothesis 3: Achievement, affiliation, personal growth, and power are negatively associated with age.

Achievement might best resemble work goals in the study of Bühler et al. (2019), which showed a strong negative association with age after age 40. Although the mean frequency of achievement-categorized goals was higher in young compared to midlife adults, Dunlop et al. (2017) did not find mean-level trends in the frequency of achievement-categorized goals within these samples. In our study, we assessed achievement broadly (i.e., beyond work-related achievement). We thus expected the achievement motive to show a slight negative age association, with no pronounced differences in older compared to younger ages. We hypothesized a negative trend in affiliation as expression of its expansive orientation

[2] The study was not preregistered. The Study of Personality Architecture and Dynamics (SPeADy) is an ongoing longitudinal research project that has been collecting data open for the scientific community to use.

towards widening the social network (Carstensen et al., 1999), and in line with findings on decreasing network size in old ages (Lang, 2000; Wrzus et al., 2013). We argue that personal growth has an expansive rather than a protective character and should thus show a negative age trend, too (compare Bühler et al., 2019). For power, pronounced age differences were likely until 60 years of age (see Bühler et al., 2019, and Dunlop et al., 2017).

Method

Participants and Procedure

Data collection of the first wave of the German research project "Study of Personality Architecture and Dynamics" (SPeADy) was performed online and via mailed questionnaires between January 2016 and January 2018. Data were collected in consistence with ethical standards for the treatment of human subjects. The project consists of two studies: the *age groups study* and the *twin family study* (Kandler, Penner, Richter, & Zapko-Willmes, 2019). In both studies, participants complete a wide range of personality-related measures. Participants of the age groups study, on which the present study was based, were recruited through various methods (e.g., distribution of flyers and posters, e-mails to mailing lists of mature-age students, and contacting sports clubs to distribute questionnaires). For the current study, only cross-sectional data of the first wave were available.[3]

Several personality-related characteristics were captured via both self- and informant reports. A total of n = 2,406 self-raters (61.6% female) answered the questions online (86.0%) or via paper-pencil version (14.0%). Mean age was 40.46 years (SD = 18.25, age range: 14–94 years). Although the sample cannot be treated as representative for the general population, it is heterogeneous regarding diverse demographics, such as family status (47.6% unwed, 41.4% married or living in a same-sex registered partnership, 2.3% permanently separated, 6.1% divorced, 2.7% widowed), educational level (0.4% left school without certificate, 4.4% currently at school, 5.9% school-leaving qualification, 16.3% secondary school certificate, 10.5% polytechnic degree, 22.1% high school graduation, 11.0% bachelor's degree, 9.5% master's degree, 14.1% state examination, 3.6% PhD, 2.0% others, and 0.2% not specified),

[3] SPeADy data of the first wave are already available as an anonymized Scientific Use File on request. Data of the second and third wave are planned to be available in April 2020 and April 2022, respectively. Researchers are welcome to contact the principal investigator Christian Kandler and to provide a short outline of their research plans (see www.speady.de/studies/?lang=en for more details).

and religion (2.0% Buddhism, 62.9% Christianity, 0.1% Hinduism, 2.4% Islam, 0.2% Judaism, 1.3% others, 31.1% none).

One male participant did not report his age, leaving a sample size of *n* = 2,405. From this sample, *n* = 2,278 participants provided ratings on their explicit motives. In a separate contact questionnaire, we asked participants to provide contact data of up to three close acquaintances. After having received the e-mail or postal address of the acquaintances, we sent them an invitation to either the online or paper-pencil version of the questionnaire together with their personalized code. The code allowed us to assign the informant reports to the respective self-report. For a total of *n* = 944 (41.4%) participants, we received one (or more) informant report(s). In our study, 95.6% of the well-informed acquaintances were spouses (29.0%), good friends (32.5%), or relatives (34.1%). Of all participants with informant reports, *n* = 224 (23.7%) participants had two, and *n* = 22 (2.3%) participants had three informant reports. If three informant reports were available, we selected two of them depending on the three following criteria (beginning with the most important one): 1. Who provided full (or more) data on the self-rater's explicit motives? 2. Who reported to be comparably closer to the self-rater? 3. Who participated first? From the *n* = 944 first raters and *n* = 246 second informant raters (total *n* = 1,190), 95.2% reported to know the self-rater very well (*n* = 781) or well (*n* = 352). Only 4.7% (*n* = 56) reported to know the self-rater moderately or less well. One person did not answer the question. *N* = 788 (66.2%) of the resulting *n* = 1,190 informant raters were females. Informant raters' mean age was 38.2 years (*SD* = 16.40, age range: 14–88 years). The data of this study can be drawn from the open science framework (*This is currently a view-only link that will be completely open after acceptance:* https://osf.io/5sepr/?view_only=25d3a8b43f94468da362389506939f31).

Measures

Intimacy, affiliation, achievement, and power motives were captured with the original German 24-item version of the Unified Motive Scales (Schönbrodt & Gerstenberg, 2012). Thirteen items were rated on a six-point Likert scale from 1 (*not important to me/the person*) to 6 (*extremely important to me/the person*). The following are item examples: "Have a close, intimate relationship with someone" (*intimacy*); "Engage in a lot of activities with other people" (*affiliation*); "Continuously engage in new, exciting, and challenging goals and projects" (*achievement*); "The opportunity to exercise control over an organization or group" (*power*; see Schönbrodt & Gerstenberg, 2012, for the German and English wording of all items). Participants rated the other eleven items on a six-point scale from 1 (*strongly*

disagree) to 6 (*strongly agree*). Informants responded to adapted third-person versions of the questions. Internal consistency coefficients ranged between .74 and .90 for the self-report, informant report 1, and informant report 2.

In addition, we used two scales of the German version of the Aspirations Index (Klusmann et al., 2005) to assess explicit motives concerning the importance of health protection and personal growth, respectively measured by five items with a four-point scale from 1 (*not important at all*) to 4 (*very important*). Loosely translated examples are "Remain untouched by diseases" (*health protection*) and "Develop my personality and learn new things" (*personal growth*; see Klusmann et al., 2005, for the German wording). Internal consistency coefficients ranged between .72 and .85 for the self-report, informant report 1, and informant report 2.

Analytical Strategy

Manifest variable analyses. Before investigating age trends in manifest item scores, we analyzed self-other agreement and inter-rater consensus. This way, we were able to ensure that within-rater correlations between constructs were not inflationary high while between-rater correlations within constructs were not. We then conducted hierarchical multiple regression analyses predicting year-by-year mean levels of explicit motives from age (linear), age^2 (quadratic), and age^3 (cubic), following the procedure by Robins, Trzesniewski, Tracy, Gosling, and Potter (2002). Since the scales have different scale ranges, we chose to *z*-standardize their scores prior to these analyses. We computed year-by-year means of explicit motives separately for females and males. This allowed us to capture aggregate age trends rather than differences between individuals within age, commonly referred to as alerting correlations. For a more detailed description of the underlying procedure, see Rosnow, Rosenthal, and Rubin (2000).

Latent variable analyses. We used a multivariate approach to test for age differences in all latent motive variables in one model. To capture mean-level age trends in explicit motives controlled for measurement error and rater biases, we analyzed motives as latent variables based on the covariance between self- and informant reports using reflexive latent factor modeling in Mplus (Muthén & Muthén, 1998–2012). Since most informants included in this study indicated to be well-acquainted with the self-rater and self-raters had no access to informants' statements, we assumed that informants provided valuable information (Huber & Power, 1985). We wanted both to ensure a good use of all information on a low level of measurement error, and to include the information given by as many informants as possible.

Therefore, we established a measurement model where one latent variable comprised three manifest variables, respectively. Those were the self-report, the informant report of acquaintance 1, and the informant report of acquaintance 2. We allowed residuals within the same rater to correlate across latent factors, thereby accounting for common variance in the manifest variables that was not accounted for by the latent variables, but by rater-specific variance (Marsh & Hau, 1996; Reuter et al., 2010). Moreover, we allowed cross-motive correlations. See Figure 1 for an illustration of the measurement and structural model.

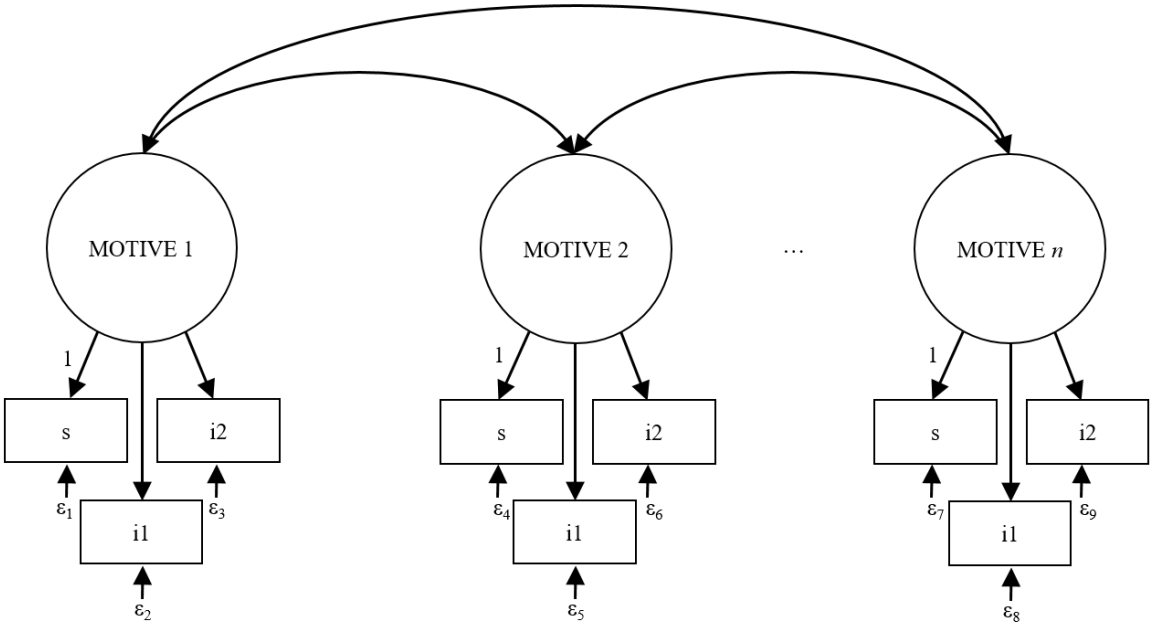

Figure 1. Measurement and structural model of explicit motives. Latent motives were captured by combining the self-report (s) with informant reports 1 (i1) and 2 (i2). Loadings of self-reports on the factor were set to 1, while loadings of informant reports were estimated freely. We included all factors in one multivariate analysis to account for shared variance. To account for rater-specific variance, we allowed residual variance of all items within each rater to correlate (not pictured for reasons of clarity). Cross-rater correlations of indicators were set to zero.

We captured age effects on explicit motives for women and men separately by including gender as a dummy-coded group variable. When comparing groups, the latent factors must have the same meaning for all of them (Muthén & Muthén, 1998–2012). To that end, we assumed and modelled strict measurement invariance by equating factor structure, factor loadings, item intercepts, and residual variances across genders (Marcusson-Clavertz & Kjell, 2018). In a multi-group design, we set factor means of males to zero (dummy control group), while factor means of females were estimated freely. We examined linear age effects as well as quadratic (age^2) and cubic (age^3) effects on latent motive variables. As Mplus works best with variances between 1 and 10 (Muthén, personal communication, September 19th, 2012), we standardized the age, age^2, and age^3 predictors as well as the manifest

indicators of the latent motive factors prior to the analysis. We stepwise tested if the predictors age³ (first step) or age² (second step) could be set to zero. This procedure is comparable to a multiple hierarchical regression in the sense that the higher-order predictor is only included if it incrementally explains variance. This way, we received the most parsimonious model. Using full information maximum likelihood, we analyzed all available information. All syntaxes can be drawn from the open science framework (*This is currently a view-only link that will be completely open after acceptance:* https://osf.io/5sepr/?view_only=25d3a8b43f94468da362389506939f31). Figure 2 shows an overview of the structural model of how we assessed age-motive associations.

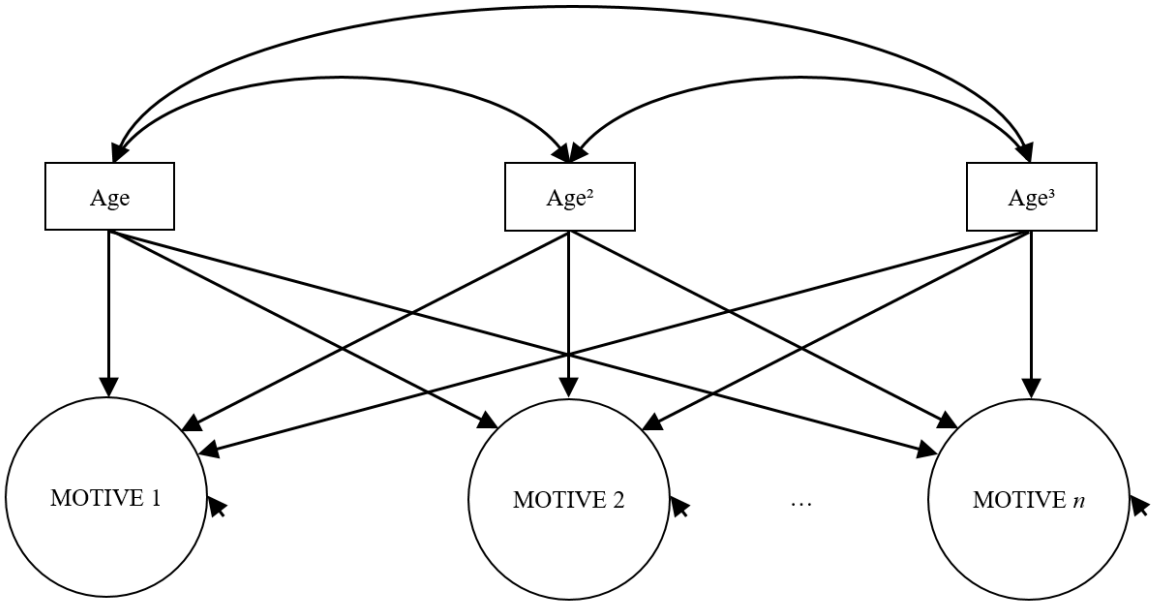

Figure 2. Structural model of age trends in latent explicit motives. We estimated the effects of age, age², and age³ separately for females and males to account for gender differences. Please note that in this figure, we printed all one-sided arrows as in the baseline model. For the results reported, however, we dropped the parameters of the third-order age effects where this did not lead to a significant model deterioration as indicated by a Wald-test of parameter constraints. We repeated this for the second-order age effects. This led to the most parsimonious model. We allowed the predictors and the dependent latent variables to correlate with each other to ensure an appropriate structural model. For reasons of clarity, cross-motive correlations are not printed.

Analytical criteria. Generally, combining RMSEA (root mean square error of approximation), CFI (comparative fit index), and SRMR (standardized root mean square residual) yields a good overall criterion of model fit for large sample sizes. Bagozzi and Yi (2012) argued that RMSEA and SRMR should be as low ($\leq$ .07) and CFI as high ($\geq$ .93) as possible. To account for the possibility of inflating Type I errors in multivariate analyses (Cribbie, 2000, 2007) and given that the large sample size making issues of general statistical power negligible (see Roberts, Caspi, & Moffitt, 2003), some authors only interpret effects

with $p < .01$ as statistically significant (compare Asendorpf, & Wilpers, 1998). Considering our smaller sample of males and possibly reduced per-parameter power by applying such an approximation (Smith & Cribbie, 2013), however, we used $p < .05$ as a criterion.

Results

Descriptive Statistics and Rater Consensus

Table 1
Mean Levels of Explicit Motives in the Total Sample and Split Per Gender

		Sample characteristics								
		Unweighted Total			Males			Females		
Motive		*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*
Protective										
Health p.	S	3.47	0.48	2,274	3.41	0.50	876	3.52	0.45	1,398
	I	3.35	0.53	932	3.29	0.57	318	3.38	0.51	614
Intimacy	S	4.61	0.77	2,275	4.47	0.82	878	4.70	0.72	1,397
	I	4.40	0.87	932	4.20	0.94	318	4.51	0.81	614
Expansive										
Achiev.	S	3.98	0.86	2,278	4.10	0.85	879	3.90	0.86	1,399
	I	3.94	0.86	932	4.03	0.93	318	3.89	0.83	614
Affiliation	S	4.07	0.96	2,276	3.85	0.96	878	4.21	0.94	1,398
	I	4.18	1.04	932	3.95	1.12	318	4.30	0.97	614
P. growth	S	3.48	0.43	2,275	3.39	0.44	877	3.54	0.41	1,398
	I	3.32	0.46	932	3.26	0.46	318	3.34	0.46	614
Power	S	3.13	1.04	2,276	3.35	1.08	878	2.99	0.99	1,398
	I	3.08	1.05	932	3.31	1.16	318	2.97	0.97	614

Note. Health p. = health protection, achiev. = achievement, p. growth = personal growth, S = self-report, I = averaged informant report.

Table 1 shows sample characteristics split by gender and rater perspective. On average and in accordance with hypothesis 1, males scored higher on achievement and power than females in both self- and informant reports, whereas females scored higher on health protection, intimacy, affiliation, and personal growth than males (see Table 2). Moreover, on average, informants tended to report lower levels of health protection, intimacy, and personal growth motives than self-raters, and had a tendency to report higher levels of the affiliation motive (see Table 2). *M* of the self-ratings were about 0.5 *SD* under the highest possible value for health protection and personal growth (see Table 1).

Table 2
Effect Sizes of Average Gender and Rater Differences

		$d_{\text{self-informant}}$			$d_{\text{female-male}}$	
Motive		Estim.	95% CI		Estim.	95% CI
Protective						
Health p.	F	0.30	0.39; 0.20	S	0.23	0.15; 0.32
	M	0.23	0.10; 0.36	I	0.17	0.03; 0.31
Intimacy	F	0.25	0.16; 0.35	S	0.30	0.22; 0.39
	M	0.32	0.19; 0.45	I	0.36	0.23; 0.50

Expansive						
Achiev.	F	0.01	–0.08; 0.11	S	–0.23	–0.32; –0.15
	M	0.08	–0.05; 0.21	I	–0.16	–0.30; –0.03
Affiliation	F	–0.10	–0.19; 0.00	S	0.38	0.30; 0.47
	M	–0.10	–0.23; 0.03	I	0.34	0.21; 0.48
P. growth	F	0.47	0.37; 0.57	S	0.36	0.27; 0.44
	M	0.29	0.16; 0.42	I	0.17	0.04; 0.31
Power	F	0.02	–0.08; 0.12	S	–0.35	–0.44; –0.27
	M	0.04	–0.09; 0.17	I	–0.33	–0.46; –0.19

Note. d = Cohen's *d*, Hedge's *g*, estim. = estimate, health p. = health protection, achiev. = achievement, p. growth = personal growth, F = female, M = male, S = self-report, I = averaged informant report.

Self-other agreement and informant rater consensus were of medium size and of comparable magnitude (see Table 3). The only exception were inter-rater correlations in personal growth, which tended to be low. Informants shared substantial variance both with each other and with the self-rater. In total, the medium correlations indicated shared variance between rater perspectives while at the same time pointing towards rater-specific differences in estimates. Supplementary Table A1 contains zero-order correlations per rater perspective across all motives. Within-rater correlations between constructs were independent of the rater perspective investigated. Moreover, on average, informant consensus and self-other agreement (Table 3) were higher than within-rater zero-order correlations (supplementary Table A1), pointing towards rater-independent construct validity of motives instead of rater-specific correlation inflations.

Table 3

Informant Consensus and Self-Other Agreement on Explicit Motives in the Total Sample

	Informant		Self-other agreement					
	consensus [a, b]		Averaged informant [c]		Informant 1 [c]		Informant 2 [b]	
Motive	*r*	95% CI	*r*	95% CI	*r*	95% CI	*r*	95% CI
Protective								
Health p.	**.40**	.28; .51	**.40**	.35; .45	**.38**	.32; .43	**.43**	.31; .53
Intimacy	**.38**	.26; .49	**.43**	.38; .48	**.41**	.36; .46	**.36**	.24; .47
Expansive								
Achiev.	**.46**	.35; .56	**.38**	.32; .43	**.36**	.30; .42	**.39**	.27; .50
Affiliation	**.56**	.46; .65	**.55**	.50; .59	**.53**	.48; .58	**.54**	.44; .63
P. growth	**.34**	.22; .45	**.26**	.20; .32	**.27**	.21; .33	.16 [d]	.03; .29
Power	**.48**	.37; .58	**.44**	.39; .49	**.43**	.38; .48	**.43**	.31; .53

Note. Health p. = health protection, achiev. = achievement, p. growth = personal growth. Correlations printed in boldface were significant at $p < .001$.

[a] Correlations across informant 1 and informant 2 ratings.

[b] N = 214-216.

[c] N = 930-932.

[d] Correlation was significant at $p = .020$.

Are there gender differences in inter-rater agreement? Self-other agreement tended to be similar for women and men. Although we found trends indicating gender differences (e.g., regarding health protection and affiliation), 95% CIs overlapped each other (see supplementary Table A2 for information on self-other agreement and informant consensus split by gender of self-raters). To back up our findings, we conducted regression analyses testing the (moderating) effect of gender on the correlation between self- and informant rating. We found no gender effects for most motives (see supplementary Table A3). However, there were main effects of gender on the informant ratings of intimacy ($b = 0.27$, $p < .001$), affiliation ($b = 0.14$, $p = .010$) and power ($b = –0.21$, $p < .001$) motives, indicating a trend of informants to rate females higher in intimacy and affiliation motives than males, and males higher in power, independently of their self-ratings. Moreover, negative interaction effects of self-rated health protection × gender ($b = –0.15$, $p = .011$) and affiliation × gender ($b = –0.14$, $p = .009$) indicated a moderating role of gender on self-other agreement in these motives. This might be interpreted as follows: For males, the predicting effect of the self-rating on the informant rating (i.e., their correlation) was higher than for females. Informant consensus tended to be similar across genders, too.

To sum up, preliminary analyses showed that rater perspectives differed regarding their mean-levels. All correlations were substantial and of similar height, in total revealing inter-rater consensus despite (random) measurement error. Besides main gender effects on informant ratings of intimacy, affiliation and power, we found interaction effects of gender on self-other agreement in health protection and affiliation.

Analyses of Manifest Age Differences Based on Self- and Informant Reports

Nonlinear hierarchical multiple regression analyses allowed a first impression on mean-level trends across different stages of life, and whether they were consistent across gender and rater perspectives. Results are presented in Tables 4 and 5. Figure 3 illustrates age trends based on the alerting mean levels per gender and rater perspective as well as based on the multivariate latent structural equation model, centered on the grand mean of age ($M = 40$, gray background). We decided to visualize results of both analytical approaches in the same figure to make their results more easily comparable.

Descriptively, rater perspectives were largely congruent regarding age trends in health protection and affiliation motives, and markedly differed regarding personal growth and males' power motives. Self- and informant reports corresponded only for certain age groups in intimacy and achievement, with reports on young males and elderly people especially

Table 4

Best-Fitting Alerting Hierarchical Regression Models for Females

		Age			Age²			Age³			Model			
											R^2 [a]		Change in F	
Motive		*b*	*SE*	*p*	*b*	*SE*	*p*	*b*	*SE*	*p*	Total	Ch. [b]	Total	$p_{ch.}$
Protective														
Health p.	S	**–0.29**	0.14	.044	**0.37**	0.12	.004				.31	.09	9.11	.004
	I	0.06	0.06	.272							.02	–	1.23	.272
Intimacy	S	**–0.24**	0.03	<.001							.45	–	59.10	<.001
	I	**2.89**	1.23	.022	**–6.57**	2.39	.008	**3.58**	1.21	.004	.19	.11	8.71	.003
Expansive														
Achiev.	S	**–0.31**	0.03	<.001							.58	–	98.59	<.001
	I	**–0.21**	0.05	<.001							.23	–	20.08	<.001
Affiliation	S	**–0.14**	0.03	<.001							.19	–	17.15	<.001
	I	–0.09	0.07	.172							.03	–	1.90	.172
P. growth	S	1.15	0.62	.067	**–2.49**	1.19	.039	**1.28**	0.59	.033	.12	.06	4.70	.033
	I	**–0.19**	0.05	.001							.15	–	12.21	.001
Power	S	**–0.22**	0.03	<.001							.43	–	55.00	<.001
	I	**–2.55**	0.91	.007	**5.36**	1.77	.003	**–2.91**	0.90	.002	.17	.13	1.52	.002

Note. N = 1,399. Age² = quadratic age, age³ = cubic age, health p. = health protection, achiev. = achievement, p. growth = personal growth, S = self-report, I = averaged informant report, *b* = unstandardized regression weight, *SE* = standard error of estimate, total = total unadjusted statistic, $p_{ch.}$ = significance of the F-test on the explained variance of the more complex model in comparison to the less complex model. Age and motive scales were standardized prior to the analyses. Significant model parameters ($p < .05$) are printed in boldface.

[a] Please note that we only decided in favor of the more complex model if the following two prerequisites both were met: The more complex model explained an incremental 5 % of the variance at minimum, and the F statistics of the more complex model reached significance ($p < .05$). This was due to the fact that there were more self-reports available than informant reports, which resulted in a higher likelihood for these statistics to reach significance.

[b] Compared to the less complex model. Please note that we did not include this estimate for regressions that included only linear age as predictor.

Table 5

Best-Fitting Alerting Hierarchical Regression Models for Males

		Age			Age²			Age³			Model			
											R^2 [a]		Change in F	
Motive		*b*	*SE*	*p*	*b*	*SE*	*p*	*b*	*SE*	*p*	Total	Ch. [b]	Total	$p_{ch.}$
Protective														
Health p.	S	–0.35	0.18	.051	**0.43**	0.15	.007				.24	.08	7.68	.007
	I	2.10	1.33	.121	–4.59	2.55	.076	**2.60**	1.28	.047	.12	.06	4.12	.047
Intimacy	S	1.34	0.71	.064	**–3.24**	1.33	.018	**1.92**	0.66	.004	.26	.09	8.62	.004
	I	**3.46**	1.39	.015	**–6.69**	2.66	.014	**3.28**	1.34	.017	.10	.09	6.02	.017
Expansive														
Achiev.	S	**2.09**	0.77	.008	**–4.67**	1.45	.002	**2.52**	0.71	.001	.21	.14	12.49	.001
	I	**1.04**	0.43	.018	**–1.00**	0.38	.010				.11	.10	6.98	.010
Affiliation	S	**–0.17**	0.05	.002							.13	–	10.42	.002
	I	–0.12	0.07	.129							.04	–	2.37	.129
P. growth	S	**–0.43**	0.19	.027	**0.36**	0.17	.034				.07	.06	4.70	.034
	I	0.00	0.08	.969							.00	–	0.00	.969
Power	S	**–0.11**	0.05	.035							.06	–	4.62	.035
	I	**0.15**	0.07	.045							.06	.06	4.19	.045

Note. N = 879. Age² = quadratic age, age³ = cubic age, health p. = health protection, achiev. = achievement, p. growth = personal growth, S = self-report, I = averaged informant report, *b* = unstandardized regression weight, *SE* = standard error of estimate, total = total unadjusted statistic, $p_{ch.}$ = significance of the *F*-test on the explained variance of the more complex model in comparison to the less complex model. Age and motive scales were standardized prior to the analyses. Significant model parameters ($p < .05$) are printed in boldface.

[a] Please note that we only decided in favor of the more complex model if the following two prerequisites both were met: The more complex model explained an incremental 5 % of the variance at minimum, and the *F* statistics of the more complex model reached significance ($p < .05$). This was due to the fact that there were more self-reports available than informant reports, which resulted in a higher likelihood for these statistics to reach significance.

[b] Compared to the less complex model. Please note that we did not include this estimate for regressions that included only linear age as predictor.

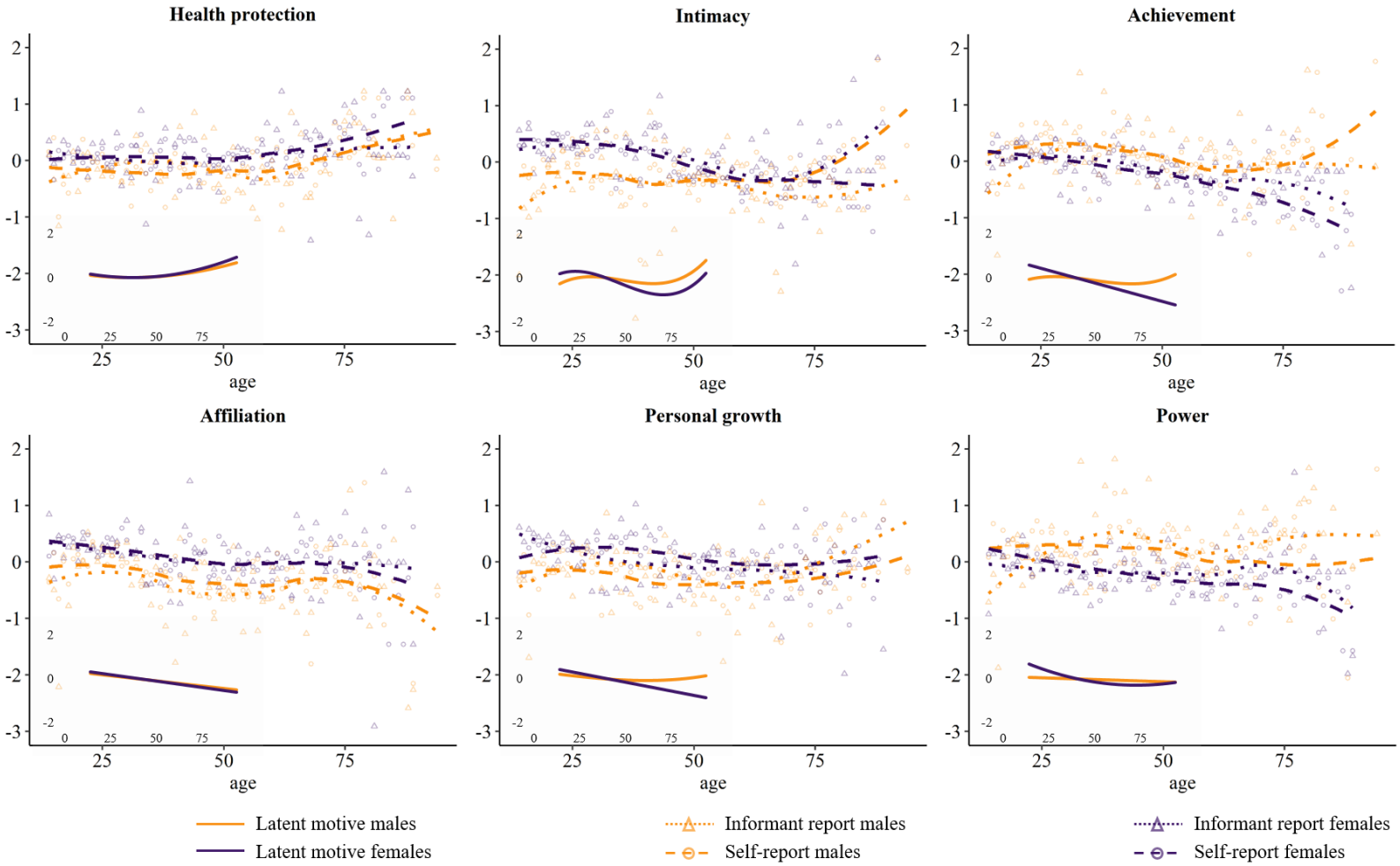

Figure 3. Alerting mean-level trends separated by gender, assessed by self-reports and averaged informant reports; and the corresponding latent mean-level trends based on self-other composite scores in explicit motives (gray background). For the manifest mean-level trends, we printed smoothed regression lines based on the standardized group means of year-by-year mean levels, which are shown as hollow shapes, using the package ggplot2 (Wickham, 2016) in R version 3.4.4 (R Core Team, 2018). The syntax of these regression lines can be drawn from the open science framework (*This is currently a view-only link that will be completely open after acceptance:* https://osf.io/5sepr/?view_only=25d3a8b43f94468da362389506939f31). For latent variable mean-level trends, we printed the unstandardized effects yielded from the most parsimonious model, and centered each group's results on the mean age of 40.

diverging. Health protection showed a positive and affiliation a negative age trend. Ratings on females' intimacy motive congruently showed a negative age trend until mature adulthood.

While females' self-rated personal growth motive followed a slight s-shaped trend, informant ratings suggested a negative association between this motive and age. Trends tended to converge for 30 to 70 years old women. Trends for males' personal growth were equally inconclusive, with a positive tendency beyond age 60. Regarding females' power motive, we found a slight negative association with age. In contrast, only males' self-rated power motive showed a negative linear age trend, while informant reports suggested a positive trend. Again, the highest divergence across rater perspectives for males' power trends appeared for males under age 30.

Does relationship duration moderate inter-rater agreement? Age trends were more comparable across raters for midlife adults than for young and elderly individuals. Given these findings, we wondered whether informant's amount and type of available information, reflected in the relationship duration of self-rater and informant, may have had an effect on self-other agreement. To the extent that informants base their ratings on information collected in a certain time frame, their ratings naturally depend on the relationship duration (i.e., time frame) and the observability of the specific motive at the current time. For instance, the relationship duration may be too short to acquire a sufficient number of cues – especially for young people who do not know each other for a long time. In contrast, informants of older people may have a substantially wider time frame on which to rely on for cues. However, certain motives that may be difficult to rate for this age group (e.g., achievement) may lead informants to exceedingly rely on their knowledge of salient behavior from the past instead of current dispositions.

To investigate the (moderating) effect of relationship duration on the correlation between self- and informant rating, we first examined the correlation between age and relationship duration ($r = .62$, $p < .001$) and then ran additional regression analyses. We found no indication of an influence of relationship duration on self-other agreement for most motives, except for personal growth (see supplementary Table A4). Independently of the self-rating, informant ratings on health protection ($b = 0.06$, $p = .048$) and power ($b = 0.09$, $p < .001$) tended to be higher, and informant ratings on intimacy ($b = –0.13$, $p < .001$) lower in longer-established compared to shorter relationships. Moreover, we found a negative interaction effect of relationship duration on the correlation between self- and informant rated

personal growth motive (b = –0.07, p = .030). That is, with increasing relationship duration, the self-other agreement in personal growth (slightly) decreased.

To sum up, self-other agreement on most motives was not predicted by relationship duration. However, rater perspectives converged more regarding personal growth for shorter than longer relationships. Moreover, informants of longer established relationships rated health protection and intimacy motives higher, but power motive lower compared to informants of more recently established relationships.

Analyses of Latent Age Differences Combining Rater Perspectives

Rater-specific age trends are not controlled for rater-specific biases and random measurement error. This might lead to unreliable estimates, especially in fringe groups. We found some, yet small evidence for moderator effects, and established that both informant consensus and self-other agreement were in an acceptable range. This makes it possible to assume that latent variables based on the shared variance of self-rating and ratings of well-informed acquaintances help to build more accurate reflections in terms of reliability and convergent validity. We thus conducted a multivariate latent variable model based on the common variance of self- and informant ratings with gender as a grouping variable. Assessing all motives in a multivariate structural equation model enabled us to control for interrelatedness (i.e., covariance) between motives, which allows more accurate conclusions based on construct-specific variance that is not confounded with a motive of similar content.

The model showed a good fit to the data with RMSEA = .017 (90%CI [.012; .021]), CFI = .980 and SRMR = .070. We first tested if the model fit significantly declined when dropping all cubic age effects that one by one had previously been indicated to not significantly differ from zero, separately for females and males. Eight of twelve possible age^3-effects were not statistically significant (see Table 6). A Wald test of parameter constraints indicated no decline in model fit [$\Delta\chi^2$= 7.40 (8), p = .494] after dropping these eight parameters. We thus reduced the full model accordingly. The resulting model fit was still good with RMSEA = .016 (90%-CI [.011; .020]), CFI = .981 and SRMR = .070. With this more parsimonious model, we repeated the procedure with the age^2-effects that one by one had been indicated not to differ from zero. Six parameters were not statistically significant (see Table 6). Again, there was no hint of deterioration in model fit after dropping these, $\Delta\chi^2$=

Table 6

Results of the Wald Tests on Setting the Effects of Age² and Age³ on Age Trends in Specific Explicit Motives to Zero

	Health protection			Intimacy			Achievement			Affiliation			Personal growth			Power		
Effects	$\Delta\chi^2$	*df*	*p*	$\Delta\chi^2$	*df*	*p*	$\Delta\chi^2$	*df*	*p*	$\Delta\chi^2$	*df*	*p*	$\Delta\chi^2$	*df*	*p*	$\Delta\chi^2$	*df*	*p*
Females																		
Age²	**8.35**	1	.004	**11.05**	1	<.001	1.27	1	.259	3.45	1	.063	0.74	1	.390	**3.91**	1	.048
Age³	0.60	1	.439	**11.17**	1	<.001	0.01	1	.905	0.01	1	.930	1.17	1	.280	0.56	1	.456
																		Males
Age²	**6.05**	1	.014	**10.28**	1	.001	**5.33**	1	.021	3.58	1	.059	2.21	1	.137	0.52	1	.471
Age³	1.76	1	.184	**13.69**	1	<.001	**8.12**	1	.004	1.06	1	.303	**5.81**	1	.016	3.25	1	.071

Note. We dropped the effects of age² and age³ on age trends in explicit motives where this did not lead to a significant model deterioration. Please note that our procedure was stepwise: We first run the tests for each effect of age³, then (after having adjusted the age³ effects where possible) of age², separately for each motive.

Significant model results ($p < .05$) are printed in boldface. Please note that we also used an α-level of 5% in this analysis.

Table 7

Model Parameter Estimates: Cross-Motive Correlations and Effects of Age, Age², and Age³ on Females' Explicit Motives as Derived From the Most Parsimonious Multivariate Analysis

Model parameters and statistics	Health protection			Intimacy			Achievement		
	Estim.	*p*	95% CI	Estim.	*p*	95% CI	Estim.	*p*	95% CI
Cross-motive correlations [a]									
Intimacy	**.16**	.032	.01; .30						
Achiev.	**.19**	.005	.06; .32	–.03	.694	–.19; .13			
Affiliation	.11	.080	–.01; .23	**.36**	<.001	.25; .47	.06	.357	–.07; .19
P. growth	**.54**	<.001	.39; .68	.10	.256	–.07; .27	**.31**	<.001	.14; .48
Power	–.02	.815	–.14; .11	.07	.285	–.06; .20	**.59**	<.001	.48; .70
Unstandardized effects									
Age →	**–0.40**	.030	–0.47; –0.03	**1.59**	.013	0.26; 1.87	**–0.41**	<.001	–0.30; –0.20
Age² →	**0.49**	.007	0.09; 0.52	**–4.27**	.001	–4.49; –1.23	0.00 [b]		0.00; 0.00
Age³ →	0.00 [b]		0.00; 0.00	**2.37**	.001	0.72; 2.46	0.00 [b]		0.00; 0.00

Note. N = 1,399. Estim. = estimate, achiev. = achievement, p. growth = personal growth. Unstandardized estimates are based on full information maximum likelihood to handle missing data. Age, age², and age³ were standardized before the analysis. Gender was dummy-coded as 0 (*male*) and 1 (*female*) and divided into two respective groups. Significant model parameters ($p < .05$) are printed in boldface.

[a] We allowed for correlations across explicit motives as they were likely to occur. This allowed us to rule out correlations of the variables of interest being uncontrolled confounds of other effects.

[b] We used a hierarchical procedure of first equating age³ effects to zero and then equating age² effects to zero, both where this was possible, to yield the most parsimonious model.

Please note that we used bootstraps (20,000) to confirm our analysis. As standard errors for bootstrapping are not reported for standardized results, significance values for the correlations are based on the model without bootstrapping.

Table 7 *continued*

Affiliation			Personal growth			Power		
Estim.	*p*	95% CI	Estim.	*p*	95% CI	Estim.	*p*	95% CI
.02	.748	–.12; .17						
.13	.026	.02; .25	.14	.082	–.02; .29			
–0.21	<.001	–0.20; –0.10	**–0.29**	<.001	–0.18; –0.09	**–0.73**	<.001	–0.72; –0.29
0.00 [b]		0.00; 0.00	0.00 [b]		0.00; 0.00	**0.46**	.003	0.11; 0.52
0.00 [b]		0.00; 0.00	0.00 [b]		0.00; 0.00	0.00 [b]		0.00; 0.00

Table 8

Model Parameter Estimates: Cross-Motive Correlations and Effects of Age, Age², and Age³ on Males' Explicit Motives as Derived From the Most Parsimonious Multivariate Analysis

Model parameters and statistics	Health protection			Intimacy			Achievement		
	Estim.	*p*	95% CI	Estim.	*p*	95% CI	Estim.	*p*	95% CI
Cross-motive correlations [a]									
Intimacy	**.27**	.002	.10; .44						
Achiev.	**.29**	<.001	.14; .44	.18	.056	–.00; .36			
Affiliation	**.26**	<.001	.12; .40	**.48**	<.001	.36; .61	**.30**	<.001	.15; .45
P. growth	**.40**	<.001	.23; .58	**.49**	<.001	.32; .66	**.63**	<.001	.46; .80
Power	.13	.085	–.02; .28	–.08	.325	–.24; .08	**.61**	<.001	.47; .75
Unstandardized effects									
Age →	–0.27	.128	–0.61; 0.08	**1.75**	.005	0.53; 2.96	0.93	.084	–0.12; 1.98
Age² →	**0.34**	.042	0.01; 0.66	**–3.80**	.003	–6.27; –1.33	–2.14	.055	–4.32; 0.04
Age³ →	0.00 [b]		0.00; 0.00	**2.04**	.003	0.72; 3.36	1.13	.063	–0.06; 2.31

Note. N = 879. Estim. = estimate, achiev. = achievement, p. growth = personal growth. Unstandardized estimates are based on full information maximum likelihood to handle missing data. Age, age², and age³ were standardized before the analysis. Gender was dummy-coded as 0 (*male*) and 1 (*female*) and divided into two respective groups. Significant model parameters ($p < .05$) are printed in boldface.

[a] We allowed for correlations across explicit motives as they were likely to occur. This allowed us to rule out correlations of the variables of interest being uncontrolled confounds of other effects.

[b] We used a hierarchical procedure of first equating age³ effects to zero and then equating age² effects to zero, both where this was possible, to yield the most parsimonious model.

Please note that we used bootstraps (20,000) to confirm our analysis. As standard errors for bootstrapping are not reported for standardized results, significance values for the correlations are based on the model without bootstrapping.

Table 8 *continued*

Affiliation			Personal growth			Power		
Estim.	*p*	95% CI	Estim.	*p*	95% CI	Estim.	*p*	95% CI
.39	<.001	.23; .56						
.27	<.001	.13; .41	**.34**	<.001	.16; .53			
–0.17	<.001	–0.23; –0.11	**–0.19**	.032	–0.36; –0.02	–0.05	.171	–0.12; 0.02
0.00 [b]		0.00; 0.00	0.00 [b]		0.00; 0.00	0.00 [b]		0.00; 0.00
0.00 [b]		0.00; 0.00	0.12	.144	–0.04; 0.29	0.00 [b]		0.00; 0.00

12.53 (6), p = .051.[4] Model fit of the resulting most parsimonious model was good with RMSEA = .017 (90%-CI [.012; .021]), CFI = .980 and SRMR = .070.

Tables 7 and 8 contain the model parameter estimates, the latent cross-motive correlations and age effects based on the most parsimonious models for females and males. As these latent trends are adjusted for measurement error and rater-specific influences, they reflect estimates based cross-rater reliability and, thus, higher convergent (cross-rater) validity. For health protection, the best fitting model indicated a nonlinear age trend, which was mainly due to a positive age association for females older than 50 years (see Table 7). The corresponding age effects were similar, but less steep for males (see Table 8). Health protection motive showed a positive age trend, in particularly beyond age 50. Intimacy followed nonlinear age trends for both genders, indicating an s-shaped curve. While the model results indicated a negative linear age effect on females' achievement motive, we did not find any significant age effects for males' achievement motive ($p < .05$). However, effects tended to be significant ($p < .10$). For both females and males, best fitting models indicated a negative linear age association for the affiliation motive (see also footnote 3 for a more detailed discussion) and the personal growth motive (see Tables 7 and 8). Females' power motive was negatively associated to age, especially for women under 50 years. Males' power motive, in contrast, did not show any significant age associations.

Based on the latent model, hypothesis 2 was especially met for the elderly, while hypothesis 3 was met for affiliation and personal growth. Achievement and power motives were negatively associated with age in females (hypothesis 3), but not in males, indicating gender differences in age-motive associations (hypothesis 1). Moreover, females' power motive showed a pronounced negative age association before age 60, supporting our specific prediction.

Discussion

Our study was designed to examine females' and males' age trends in explicit motives based on cross-sectional data of self- and informant reports in a German sample across the

[4] We acknowledge that χ^2-test statistics tended to become significant, specifically in the case of quadratic age effects on the affiliation motive (see Table 6). Equating age² effects on affiliation for both genders at the same time after having equated all age³ and all other age² effects according to our procedure indeed yielded a significant model deterioration, $\Delta\chi^2$= 7.38 (2), p = .025. The resulting age trends, with estimated age² effects on affiliation, can be found in supplementary Figure A1. None of the specific effects (b_{AGE^2} = .24 and b_{AGE^2} = .33), however, reached statistical significance (p = .057 and p = .054).

ages 14–94 years. The inclusion of gender and the assessments from well-acquainted informants were methodological novelties of our study in comparison to previous investigations in this field. Informant reports allowed to compare rater perspectives on age-motive associations, and yielded the possibility of more reliable reflections of the measured constructs compared to self-ratings only. Except for the study by Bühler et al. (2019), previous studies on age-goal associations have in common that they were based on frequency ratings of goals in certain life domains. Differently from this approach, we used dimensional importance ratings on motives, which are more standardized than an open format. Most motives showed age associations pointing towards predictions of lifespan theories, especially in older ages. However, we did not find age trends for males' achievement and power motives. Before we discuss these findings, we will reflect on the findings on rater convergence in this study.

Rater Convergence

In sum, rater convergence was acceptable for the total sample. Correlations were of medium magnitude, and informant consensus and self-other agreement did not substantially differ between each other. This implies three things: First, findings of self-other agreement yielded shared variance between rater perspectives, implying that not only the self-rater can evaluate their motives, but informant raters also have access to them. Second, rater-specific variance seems to be independent of the observer perspective (i.e., self-raters and informant raters did not differ regarding the magnitude of rater-specific variance). Third, informant consensus being of the same magnitude as self-other agreement means that there is consensus between observers that indicates a common perspective, and might even explain variance that is not accessible to or reported by the self-rater.

However, that correlations were medium also implied substantial rater-specific variance. Besides measurement error, this could point to rater-specific biases in motive evaluations. In fact, across all age groups, self-other agreement on health protection and affiliation tended to be higher for males than for females. Moreover, self- and informant rated age trends corresponded more for midlife adults than for very young and elderly individuals. At the same time, we received the smallest response rates for people under 20 years, between 30 to 45 years, and people over 60 years. For the youngest cohort, we recruited a very small age range (14–20 years), naturally leading to a comparably small group. Similarly, few elderly participants might explain in part why self-other agreement for very young and very old participants was lower than for other ages. However, this was not true for the ages 30–45

years where we received some of the highest agreement despite some of the lowest response rates. Supplemental analyses yielded that relationship duration could not explain why older people's self- and informant ratings tended to diverge. The only exception were small negative effects of relationship duration on self-other agreement on personal growth ratings. This might indicate that long-term acquaintances rely their evaluations on self-raters' personal growth on *prior* instead of actual behavior. Moreover, personal growth motive tended to show the lowest inter-rater agreement. This might either be a hint that informants are generally not well-equipped to rate the personal-growth motive of their significant others. It could also imply that self-raters tend to deceive themselves and/or researchers when reporting on them. The latter perspective is slightly supported by high mean levels of self-ratings in this motive.

Another reason for higher rater overlaps in midlife adults' compared to other age groups' age trends might be that midlife adults can articulate and act more in consensus with their motives. Very young people, in contrast, cannot yet fully choose their environment (for example, in Germany, there is compulsory school attendance until the age of 18). This might make it harder for youngsters to act in accordance with their motives, and in turn makes it harder for observers to assume them (although youngsters can of course articulate their motives to significant others). Moreover, the years of the late teens to the late 20s are characterized by a high instability in roles and commitment (Arnett, 2000), possibly also indicating uncertainty regarding motives. Put differently, lower self-other agreement might not only be attributable to lower informant accuracy, but might instead be a problem of reduced introspection of younger people.

For older people, it might not be as easy (anymore) to live in accordance with their motives, either. For example, there might be an elderly woman with a high need for intimacy who, unfortunately, has already lost her spouse and is not so steady on her feet. Not seeing her leave the house might let neighbors assume that she is content with being alone, while in truth, she is physically not able to behave in accordance with her intimacy motive. Interestingly, while elderly females' self-ratings indicated higher levels in intimacy and achievement than reported by their well-informed acquaintances, the opposite was true for elderly males. For example, it might be true that an elderly man feels a moderate need for intimacy, while his children think he must be very lonely, and just does not show it. Of course, these examples are highly pithy and moreover very speculative based on our data at hand. To conclude, it might be the case that acquaintances relied their ratings on actual behavior or observable positions, while self-reports were more oriented towards individuals'

feelings and cognitive reflections. We could, however, also imagine that either self-deception or impression-management tendencies have led to some of the divergences.

Can Lifespan Theories Explain Age Trends in Explicit Motives?

In general, our findings supported lifespan theories insofar as we found positive age trends in protective motives for people of older ages and negative age trends in two of the four proposed expansive motives. However, our results revealed a more complex pattern for younger age groups and between genders.

Age trends in protective motives. We found positive age trends for health protection and intimacy, which were pronounced for older people. In contrast to our findings, Heckhausen (1997) reported an increased importance of health for midlife adults (40–55 years old) in comparison to young adults. One reason for this discrepancy might be the increasing life expectancy in Germany (Bundesinstitut für Bevölkerungsforschung, 2018), leading to decreased awareness of mortality (and affecting the subjective perception of time) in today's society. This would explain why in our sample, adults did not show positive age trends in their health protection motive before the age of 55, while 20 years ago, this shift was observable for comparably younger individuals. However, different operationalizations of health motivation might also explain part of the difference.

We found a positive intimacy-age association for emerging adults, which may reflect their increasing orientation towards a lasting romantic relationship (compare Arnett, 2000). Many people in Germany still look for and start a family until their early 30s. For example, in 2018, females' mean age at marriage in Germany was 32.1 years, while males' mean age was 34.6 years (Statistisches Bundesamt, 2019). Given that the items of intimacy were mostly related to a close romantic relationship, higher intimacy levels in younger compared to older adults might not only reflect more motivation to protect present relationships, but also to invest in (and build) close relationships (e.g., partner, family). The negative trend in intimacy between 30 to 60 years corresponds to Heckhausen's (1997) finding of fewer reports of developmental goals regarding family. According to Brandtstädter et al. (2010), the reason for specific age trends in motives also lays in the fact that individuals strive to contain a positive self-view, which can only be upheld if they adapt motives to be realistic, or, put differently, controllable (see also Heckhausen, 1997). In line with this, a positive age association with motive levels in a certain time span with a negative age association thereafter (here, after age 35) can be a hint on developmental deadlines (Heckhausen, Wrosch, & Fleeson, 2001). In this case, one might speculate that younger women and men seek a deep relationship with the

possibility to establish a family, while midlife individuals either have already got their families or have disengaged from this goal. However, the opposite could also be true: Midlife adults might have changed priorities due to separation or children having left home. These assumptions cannot be drawn solely based on our data. However, they might be an interesting starting point for future investigations.

The age of 65 years marks another turning point of adult people's life: Approaching retirement, which implies inhabiting a new social role. According to the social investment theory, taking over (or leaving) normative (e.g., family- or work-related) roles is associated with changes in personality traits (Hudson, Roberts, & Lodi-Smith, 2012). Applying this theory to positive age trends in the elderly's mean intimacy motive, old age might not only reflect more motivation to protect present relationships, but might partly come along with a normative work-related change, which has an influence on adults' social roles and time management. It might also be true that perceiving life as finite within a more or less predictable time span itself, which might be triggered by retirement, leads to a more pronounced orientation towards significant others (Carstensen et al., 1999).

Age trends in expansive motives. Whereas we found the expected negative age trends in affiliation and personal growth across genders, our analyses revealed considerable gender differences in age trends in achievement and power. Specifically, we could only confirm the hypothesized age trends for females.

One explanation for males' (missing) achievement trends might be that job-related goal-attainment must often be reached within a limited time frame (see, for example, Heckhausen, 1997, on developmental deadlines in career paths), and that male adults might invest more in this outcome directly prior to the window closing, i.e., in midlife. Thus, a general negative age trend might have been confounded with the closing window of opportunity. However, besides their closeness to some findings on implicit motives (Veroff, Reuman, & Feld, 1984; Valero, Nikitin, & Freund, 2014), the non-significant age trends in males' achievement could also reflect differences in cohort-related achievement motives (compare Dobewall, Tormos, & Vauclair, 2017). For example, among older generations in Germany, men were (primarily) the ones who worked and pursued a career, whereas women often took care of family and household. This might have resulted in a non-observable age trend in males' achievement, where a negative age trend combines with a positive cohort effect (that is, the older the cohort, the stronger the motive). Among younger generations, however, this trend might reverse as females see their new-gained chances to attain major

career outcomes. Thus, it might be possible that a negative cohort effect for females (that is, the younger the cohort, the higher the motive) drove part of these gender differences in age trends. Contrary to this perspective, Bühler et al. (2019) found a strong negative association of work goals with age. In line with this, the conceptualization of the achievement motive in our study was not only reduced to work, but could be applied to any context in life (e.g., sports). Possibly, older men simply have a higher motive to do generally well in life compared to older women, which is not reduced to a certain life domain.

For power, a cohort effect might be possible, too. Nowadays, females might become more power-oriented, while males might become less power-oriented along with changes in gender-related traditional roles in Germany and other Western societies. This trend would be able to cover up actual changes in power within individuals as found by Dunlop et al. (2017), and could explain the exact picture that we found in our study: No age trends for males with a steep negative age trend for younger females. A comparison of self- and informant rated power trends revealed another interesting finding: While self-rated power showed a negative age trend, informants reported a positive trend in power until about 50 years of age. Males' self-rated motive to gain power might not change or is negatively associated with age, while in view of others, males' power shows a positive trend until midlife, possibly as part of perceived job-related engagement and achievements (compare Heckhausen, 1997; see Veroff et al., 1984, on higher implicit hope for power in midlife males). Males might feel that their *need* for power stays stable (or shows a negative trend) across age, captured by self-reports, while (or *because*) their observable power levels show a positive age trend, interpreted by informants as a higher power motive. However, this interpretation is highly speculative and should thus be considered with caution.

Limitations and Future Directions

Although we used a large sample of more than 2,000 self-raters and about 1,000 informants between adolescence and old age and considered convergence between self- and informant reports to yield more accurate assessments of our participants' motives, our study still has limitations. The most important limitation is our cross-sectional design. When examining age trends, it is best to use a sample of different cohorts and measure their traits multiple times to be able to disentangle age from cohort effects. With the current data at hand, we could only model cross-sectional age differences. As outlined above, we cannot simply draw the conclusion that male individuals in general do not show age trends in achievement and power. It is just as likely that these findings represent trends in motives that are

attributable to gender differences in both age- and cohort-related shifts. The study by Dobewall et al. (2017) has acknowledged the importance of considering both age and cohort effects in particular in face of generational shifts in role models. Future research should address this issue more deeply regarding motives by analyzing longitudinal data including multiple cohorts.

Second, we had a rare number of participants in some ages, especially for men. This heightened standard errors of measurement and might have skewed our estimates. The latter may account for the higher rater congruence in females' compared to males' age trends due to a more accurate (i.e., reliable) reflection of mean estimates. Moreover, lower reliability in males' estimates might also have covered up possibly true age effects on males' achievement. That is, age trends in males' achievement did not reach statistical significance, although effect sizes did not differ markedly between genders.

Moreover, supplemental analyses showed that gender and relationship duration might moderate self-other agreement on explicit motives. Although these effects were small, they pointed towards higher self-other agreement in males' compared to females' health protection and affiliation motives, and towards higher self-other agreement on personal growth in shorter compared to longer relationships. It might be interesting to investigate gender differences in self-other agreement and informant consensus on motives as well as other possible moderators more deeply in the future to explain why these differences occurred.

Although the possibility to examine multiple latent associations is a major advantage in structural equation models, simultaneously testing multiple parameters can inflate the Type I error rate (Cribbie, 2000, 2007). Although the number of participants in the present study was sufficiently high to neglect issues of *general* statistical power (compare Roberts et al., 2003), research has shown that correction procedures can lead to overly conservative tests and reduced (per-parameter) power (Smith & Cribbie, 2013). We want to stress that a more conservative procedure would have enhanced the risk for low per-parameter power and, thus, Type II errors (i.e., the risk to neglect finding associations that might in reality have been there). Thus, we decided to stick with an alpha level of $p < .05$.

Finally, we only asked participants that were German speakers, mainly located in Germany, Austria, and Switzerland. Although age-related trends in value priorities are generalizable across cultures (Fung, Ho, Zhang, Zhang, Noels, & Tam, 2016), it is well-known that life concepts in industrialized, individualized societies differ from those in communal societies. This may come along with different age- and gender-related trends in

motives. Thus, future studies should examine multiple populations of different cultures and languages to proof generalization of our results.

Conclusions

Our study served as one of the first to address gender-specific age trends in explicit motives across a very broad age range with the use of both self- and informant reports. This allowed more reliable information on age trends in individual motives across the lifespan. Lifespan theories might serve as theoretical frameworks to partially explain age-related shifts in explicit motives. For example, health protection showed a positive age association for individuals over 50 years. Moreover, age-graded windows of opportunity might shape age-related shifts in motives as might be concluded, for instance, from the reversed s-shaped intimacy-age association across the lifespan. However, we also showed that both age and gender play a role in shaping explicit motives across adulthood. Normative gender roles have changed during the last 50 years. Our findings might in part imply that age trends in motives may represent cohort differences and zeitgeist. For men, this was especially salient in non-significant achievement and power trends. For women under age 30, negative power-age associations stress this conclusion. Future studies should prove whether our findings can be replicated with large multiple-wave multiple-cohort studies using multiple informants to disentangle age-related developmental trends from cohort effects.

References

Arnett, J. J. (2000). Emerging adulthood: A theory of development from the late teens through the twenties. *American Psychologist, 55*, 469–485. https://doi.org/10.1037//0003-066X.55.5.469

Asendorpf, J. B., & Wilpers, S. (1998). Personality effects on social relationships. *Journal of Personality and Social Psychology, 74*, 1531–1544. https://doi.org/10.1037/0022-3514.74.6.1531

Bagozzi, R. P., & Yi, Y. (2012). Specification, evaluation, and interpretation of structural equation models. *Journal of the Academy of Marketing Science, 40*, 8–34. https://doi.org/10.1007/s11747-011-0278-x

Baltes, P. B., & Baltes, M. M. (1990). Psychological perspectives on successful aging: The model of selective optimization with compensation. In M. M. Baltes & P. B. Baltes (Eds.), *European network on longitudinal studies on individual development. Successful aging: Perspectives from the behavioral sciences* (pp. 1–34). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511665684.003

Bleidorn, W., Kandler, C., Hülsheger, U. R., Riemann, R., Angleitner, A., & Spinath, F. M. (2010). Nature and nurture of the interplay between personality traits and major life goals. *Journal of Personality and Social Psychology, 99*, 366–379. https://doi.org/10.1037/a0019982

Borg, I., Hertel, G., & Hermann, D. (2017). Age and personal values: Similar value circles with shifting priorities. *Psychology and Aging, 32*, 636–641. https://doi.org/10.1037/pag0000196

Brandtstädter, J., Rothermund, K., Kranz, D., & Kühn, W. (2010). Final decentrations. *European Psychologist, 15*, 152–163. https://doi.org/10.1027/1016-9040/a000019

Brandtstädter, J., Wentura, D., & Greve, W. (1993). Adaptive resources of the aging self: Outlines of an emergent perspective. *International Journal of Behavioral Development, 16*, 323–349. https://doi.org/10.1177/016502549301600212

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

Bundesinstitut für Bevölkerungsforschung (2018). Lebenserwartung bei Geburt nach Geschlecht, Sterbetafel 1871/1881 bis 2014/2016 (Life expectancy at birth per gender, mortality tables 1871/1881 to 2014/2016) [Website]. Retrieved from https://www.bib.bund.de/DE/Fakten/Fakt/S31-Lebenserwartung-Geburt-Geschlecht-ab-1871.html?nn=9992060

Carstensen, L. L. (2006). The influence of a sense of time on human development. *Science (New York, N.Y.), 312*, 1913–1915. https://doi.org/10.1126/science.1127488

Carstensen, L. L., & Fredrickson, B. L. (1998). Influence of HIV status and age on cognitive representations of others. *Health Psychology, 17*, 494–503. https://doi.org/10.1037/0278-6133.17.6.494

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

Cribbie, R. A. (2000). Evaluating the importance of individual parameters in structural equation modeling: The need for Type I error control. *Personality and Individual Differences*, *29*, 567–577. https://doi.org/10.1016/S0191-8869(99)00219-6

Cribbie, R. A. (2007). Multiplicity control in structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *14*, 98–112. https://doi.org/10.1207/s15328007sem1401_5

Davidov, E., Schmidt, P., & Schwartz, S. H. (2008). Bringing values back in: The adequacy of the European Social Survey to measure values in 20 countries. *Public Opinion Quarterly*, *72*, 420–445. https://doi.org/10.1093/poq/nfn035

Denzinger, F., Backes, S., Job, V., & Brandstätter, V. (2016). Age and gender differences in implicit motives. *Journal of Research in Personality*, *65*, 52–61. https://doi.org/10.1016/j.jrp.2016.09.003

Dobewall, H., & Aavik, T. (2016). Rank-order consistency and profile stability of self- and informant-reports of personal values in comparison to personality traits. *Journal of Individual Differences*, *37*, 40–48. https://doi.org/10.1027/1614-0001/a000186

Dobewall, H., Tormos, R., & Vauclair, C.-M. (2017). Normative value change across the human life cycle: Similarities and differences across Europe. *Journal of Adult Development*, *24*, 263–276. https://doi.org/10.1007/s10804-017-9264-y

Dunlop, W. L., Bannon, B. L., & McAdams, D. P. (2017). Studying the motivated agent through time: Personal goal development during the adult life span. *Journal of Personality*, *85*, 207–219. https://doi.org/10.1111/jopy.12234

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

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

Freund, A. M. (2006). Age-differential motivational consequences of optimization versus compensation focus in younger and older adults. *Psychology and Aging, 21*, 240–252. https://doi.org/10.1037/0882-7974.21.2.240

Freund, A. M., & Ebner, N. C. (2005). The aging self: Shifting from promoting gains to balancing losses. In W. Greve, K. Rothermund, & D. Wentura (Eds.), *The adaptive self: Personal continuity and intentional self-development* (pp. 185–202). Ashland, OH: Hogrefe & Huber. Retrieved from http://web.a.ebscohost.com/ehost/detail/detail?vid=8&sid=ce804550-7896-4438-a4c5-02b3aa1d5929%40sessionmgr4007&bdata=JnNpdGU9ZWhvc3QtbGl2ZQ%3d%3d#AN=2005-02582-010&db=psyh

Freund, A. M., Hennecke, M., & Riedinger, M. (2008). Age-related differences in outcome and process goal focus. *European Journal of Developmental Psychology, 7*, 198–222. https://doi.org/10.1080/17405620801969585

Funder, D. C., & Sneed, C. D. (1993). Behavioral manifestations of personality: An ecological approach to judgmental accuracy. *Journal of Personality and Social Psychology, 64*, 479–490. https://doi.org/10.1037/0022-3514.64.3.479

Fung, H. H., Ho, Y. W., Zhang, R., Zhang, X., Noels, K. A., & Tam, K.-P. (2016). Age differences in personal values: Universal or cultural specific? *Psychology and Aging, 31*, 274–286. https://doi.org/10.1037/pag0000082

Hagemeyer, B., Dufner, M., & Denissen, J. J. A. (2016). Double dissociation between implicit and explicit affiliative motives: A closer look at socializing behavior in dyadic interactions. *Journal of Research in Personality, 65*, 89–93. https://doi.org/10.1016/j.jrp.2016.08.003

Heckhausen, J. (1997). Developmental regulation across adulthood: Primary and secondary control of age-related challenges. *Developmental Psychology, 33*, 176–187. https://doi.org/10.1037/0012-1649.33.1.176

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

Heckhausen, J., & Schulz, R. (1995). A life-span theory of control. *Psychological Review, 102*, 284–304. https://doi.org/10.1037/0033-295X.102.2.284

Heckhausen, J., Wrosch, C., & Fleeson, W. (2001). Developmental regulation before and after a developmental deadline: The sample case of "biological clock" for childbearing. *Psychology and Aging, 16*, 400–413. https://doi.org/10.1037/0882-7974.16.3.400

Heckhausen, J., Wrosch, C., & Schulz, R. (2010). A motivational theory of life-span development. *Psychological Review, 117*, 32–60. https://doi.org/10.1037/a0017668

Hofer, J., Busch, H., Bond, M. H., Li, M., & Law, R. (2010). Effects of motive-goal congruence on well-being in the power domain: Considering goals and values in a German and two Chinese samples. *Journal of Research in Personality, 44*, 610–620. https://doi.org/10.1016/j.jrp.2010.08.001

Huber, G. P., & Power, D. J. (1985). Retrospective reports of strategic-level managers: Guidelines for increasing their accuracy. *Strategic Management Journal, 6*, 171–180.

Hudson, N. W., Roberts, B. W., & Lodi-Smith, J. (2012). Personality trait development and social investment in work. *Journal of Research in Personality, 46*, 334–344. https://doi.org/10.1016/j.jrp.2012.03.002

Kandler, C., Penner, A., Richter, J., & Zapko-Willmes, A. (2019). The Study of Personality Architecture and Dynamics (SPeADy): A longitudinal and extended twin family study. *Twin Research and Human Genetics, 22(6)*, 1-6. https://doi.org/10.1017/thg.2019.62

Kasser, T., & Ryan, R. M. (1996). Further examining the American dream: Differential correlates of intrinsic and extrinsic goals. *Personality and Social Psychology Bulletin, 22*, 280–287. https://doi.org/10.1177/0146167296223006

Kim, H., Di Domenico, S. I., & Connelly, B. S. (2018). Self–other agreement in personality reports: A meta-analytic comparison of self- and informant-report means. *Psychological Science,* 1–10, https://doi.org/10.1177/0956797618810000

Klusmann, U., Trautwein, U., & Lüdtke, O. (2005). Intrinsische und extrinsische Lebensziele: Reliabilität und Validität einer deutschen Fassung des Aspirations Index [Intrinsic and extrinsic life goals: Reliability and validity of a German version of the Aspirations Index]. *Diagnostica, 51*, 40–51. https://doi.org/10.1026/0012-1924.51.1.40

Köllner, M. G., & Schultheiss, O. C. (2014). Meta-analytic evidence of low convergence between implicit and explicit measures of the needs for achievement, affiliation, and power. *Frontiers in Psychology, 5(826)*, 1–30. doi:10.3389/fpsyg.2014.00826

Lang, F. R. (2000). Ending and continuity of social relationships: Maximizing intrinsic benefits within personal networks when feeling near to death. *Journal of Social and Personal Relationships, 17*, 155–182. https://doi.org/10.1177/0265407500172001

Letzring, T. D., Wells, S. M., & Funder, D. C. (2006). Information quantity and quality affect the realistic accuracy of personality judgment. *Journal of Personality and Social Psychology, 91*, 111–123. https://doi.org/10.1037/0022-3514.91.1.111

Marcusson-Clavertz, D., & Kjell, O. N. E. (2018). Psychometric properties of the Spontaneous and Deliberate Mind Wandering scales. *European Journal of Psychological Assessment,* 1–13. https://doi.org/10.1027/1015-5759/a000470

Marsh, H. W., & Hau, K.-T. (1996). Assessing goodness of fit: Is parsimony always desirable? *The Journal of Experimental Education*, *64*, 364–390. https://doi.org/10.1080/00220973.1996.10806604

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

McClelland, D. C., Koestner, R., & Weinberger, J. (1989). How do self-attributed and implicit motives differ? *Psychological Review, 96*, 690–702. doi:10.1037/0033-295X.96.4.690

Mroczek, D. K. (2001). Age and emotion in adulthood. *Current Directions in Psychological Science, 10*(3), 87–90. https://doi.org/10.1111/1467-8721.00122

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

Muthén, L. K. (2012, September 19). Re: No convergence. # of iterations exceeded [Online forum comment]. Retrieved from http://www.statmodel.com/discussion/messages/9/7079.html?1452653862

Muthén, L. K., & Muthén, B. O. (1998–2012). *Mplus User's Guide: Seventh Edition.* Los Angeles, CA: Muthén & Muthén. Retrieved from https://www.statmodel.com/download/usersguide/Mplus%20user%20guide%20Ver_7_r3_web.pdf

Nurmi, J.-E. (1992). Age differences in adult life goals, concerns, and their temporal extension: A life course approach to future-oriented motivation. *International Journal of Behavioral Development, 15*, 487–508. https://doi.org/10.1177/016502549201500404

R Core Team (2018). R: A language and environment for statistical computing (Version 3.4.4) [Programming language]. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from https://www.R-project.org/

Reuter, T., Ziegelmann, J. P., Wiedemann, A. U., Geiser, C., Lippke, S., Schüz, B., & Schwarzer, R. (2010). Changes in intentions, planning, and self-efficacy predict changes in behaviors: An application of latent true change modeling. *Journal of Health Psychology*, *15*, 935–947. https://doi.org/10.1177/1359105309360071

Ritter, J., & Freund, A. M. (2014). Values across adulthood: A neglected developmental construct. In P. Verhaeghen & C. K. Hertzog (Eds.), *The Oxford handbook of emotion, social cognition, and problem solving in adulthood.* New York: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199899463.013.017

Roberts, B. W., Caspi, A., & Moffitt, T. E. (2003). Work experiences and personality development in young adulthood. *Journal of Personality and Social Psychology*, *84*, 582–593. https://doi.org/10.1037/0022-3514.84.3.582

Roberts, B. W., O'Donnell, M., & Robins, R. W. (2004). Goal and personality trait development in emerging adulthood. *Journal of Personality and Social Psychology, 87,* 541–550. https://doi.org/10.1037/0022-3514.87.4.541

Roberts, B. W., & Robins, R. W. (2000). Broad dispositions, broad aspirations: The intersection of personality traits and major life goals. *Personality and Social Psychology Bulletin*, *26*, 1284–1296. https://doi.org/10.1177/0146167200262009

Robins, R. W., Trzesniewski, K. H., Tracy, J. L., Gosling, S. D., & Potter, J. (2002). Global self-esteem across the life span. *Psychology and Aging, 17*, 423–434. https://doi.org/10.1037/0882-7974.17.3.423

Robinson, O. C. (2013). Values and adult age: Findings from two cohorts of the European Social Survey. *European Journal of Ageing, 10,* 11–23. https://doi.org/10.1007/s10433-012-0247-3

Rosnow, R. L., Rosenthal, R., & Rubin, D. B. (2000). Contrasts and correlations in effect-size estimation. Psychological Science, 11, 446–453. https://doi.org/10.1111/1467-9280.00287

Rothermund, K., & Brandtstädter, J. (2003). Coping with deficits and losses in later life: From compensatory action to accommodation. *Psychology and Aging, 18*, 896–905. https://doi.org/10.1037/0882-7974.18.4.896

Schönbrodt, F. D., & Gerstenberg, F. X. R. (2012). An IRT analysis of motive questionnaires: The Unified Motive Scales. *Journal of Research in Personality, 46*, 725–742. https://doi.org/10.1016/j.jrp.2012.08.010

Schultheiss, O. C., & Brunstein, J. C. (2001). Goal imagery: Bridging the gap between implicit motives and explicit goals. *Journal of Personality, 67*, 1–38, https://doi.org/10.1111/1467-6494.00046

Schwartz, S. H., Cieciuch, J., Vecchione, M., Davidov, E., Fischer, R., Beierlein, C., . . . Konty, M. (2012). Refining the theory of basic individual values. *Journal of Personality and Social Psychology, 103*, 663–688. https://doi.org/10.1037/a0029393

Sheldon, K. M., & Cooper, M. L. (2008). Goal striving within agentic and communal roles: Separate but functionally similar pathways to enhanced well-being. *Journal of Personality, 76*, 415–447. https://doi.org/10.1111/j.1467-6494.2008.00491.x

Smith, C. E., & Cribbie, R. A. (2013). Multiplicity control in structural equation modeling: Incorporating parameter dependencies. *Structural Equation Modeling: A Multidisciplinary Journal, 20*, 79–85. https://doi.org/10.1080/10705511.2013.742385

Soto, C. J., John, O. P., Gosling, S. D., & Potter, J. (2011). Age differences in personality traits from 10 to 65: Big Five domains and facets in a large cross-sectional sample. *Journal of Personality and Social Psychology, 100*, 330–348. https://doi.org/10.1037/a0021717

Statistisches Bundesamt (2019). Eheschließungen, Ehescheidungen und Lebenspartnerschaften – Daten zu den Eheschließungen und dem durchschnittlichen Heiratsalter Lediger (Marriages, divorces, and regristered partnerships – Data on marriages and the average age to enter marriage) [Website]. Retrieved from https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bevoelkerung/Eheschliessungen-Ehescheidungen-Lebenspartnerschaften/Tabellen/eheschliessungen-heiratsalter.html

Vaillant, G. E. (2007). Aging well. *The American Journal of Geriatric Psychiatry, 15*, 181–183. https://doi.org/10.1097/JGP.0b013e31803190e0

Valero, D., Nikitin, J., & Freund, A. (2014). The effect of age and time perspective on implicit motives. *Motivation and Emotion, 2*, 175–181. https://doi.org/10.1007/s11031-014-9453-y

Veroff, J., Reuman, D., & Feld, S. (1984). Motives in American men and women across the adult life span. *Developmental Psychology, 20*(6), 1142–1158. https://doi.org/10.1037/0012-1649.20.6.1142

Wagerman, S. A., & Funder, D. C. (2007). Acquaintance reports of personality and academic achievement: A case for conscientiousness. *Journal of Research in Personality, 41*, 221–229. https://doi.org/10.1016/j.jrp.2006.03.001

Wagner, S. M., Rau, C., & Lindemann, E. (2010). Multiple informant methodology: A critical review and recommendations. *Sociological Methods & Research, 38*, 582–618. https://doi.org/10.1177/0049124110366231

Wickham, H. (2016). ggplot2: Elegant graphics for data analysis. Springer Publishing, New York. Retrieved from https://ggplot2.tidyverse.org

Wrzus, C., Hänel, M., Wagner, J., & Neyer, F. J. (2013). Social network changes and life events across the life span: A meta-analysis. *Psychological Bulletin, 139,* 53–80. https://doi.org/10.1037/a0028601

Online Supplement: Tables and Figures

Table A1

Zero-Order Correlations per Rater Perspective

	Self [a]\averaged informant [b]						Informant 1 [c]\informant 2 [d]					
Motive	Health p.	Intimacy	Achiev.	Affil.	P. growth	Power	Health p.	Intimacy	Achiev.	Affil.	P. growth	Power
Health p.		**.18**	**.22**	**.22**	**.49**	.04		**.25**	**.21**	**.23**	**.53**	–.05
Intimacy	**.24**		**.15**	**.41**	**.29**	**–.07**	**.19**		**.18**	**.38**	**.30**	–.01
Achiev.	**.19**	**.22**		**.18**	**.42**	**.47**	**.25**	**.16**		**.18**	**.44**	**.62**
Affil.	**.21**	**.44**	**.22**		**.29**	.06	**.22**	**.40**	**.19**		**.19**	.12
P. growth	**.50**	**.31**	**.30**	**.25**		.08	**.49**	**.29**	**.42**	**.30**		**.19**
Power	.05	.05	**.49**	**.15**	**.12**		.07	–.07	**.45**	.07	.07	

Note. Health p. = health protection, achiev. = achievement, affil. = affiliation, p. growth = personal growth. Correlations printed in boldface were significant at $p < .01$.

[a] N = 2,272-2,276; [b] n = 932; [c] n = 931; [d] n = 215.

Table A2

Gender-Specific Informant Consensus and Self-Other Agreement on Explicit Motives

		Informant consensus [a, b]		Self-other agreement					
				Averaged informant [c]		Informant 1 [c]		Informant 2 [b]	
Motive		*r*	95% CI	*r*	95% CI	*r*	95% CI	*r*	95% CI
Protective									
Health p.	F	**.45**	.31; .57	**.34**	.27; .41	**.33**	.26; .40	**.33**	.18; .46
	M	.31 [d]	.07; .52	**.48**	.39; .56	**.44**	.35; .53	**.62**	.44; .75
Intimacy	F	**.41**	.27; .53	**.42**	.35; .48	**.40**	.33; .46	**.31**	.16; .45
	M	.32 [d]	.08; .53	**.43**	.34; .52	**.40**	.30; .49	**.44**	.21; .62
Expansive									
Achiev.	F	**.43**	.29; .55	**.39**	.32; .46	**.36**	.29; .43	**.42**	.28; .54
	M	**.49**	.27; .66	**.36**	.26; .45	**.36**	.26; .45	.29 [d]	.04; .50
Affiliation	F	**.49**	.36; .60	**.51**	.45; .57	**.49**	.43; .55	**.45**	.31; .57
	M	**.63**	.45; .76	**.57**	.49; .64	**.56**	.48; .63	**.68**	.52; .80
P. growth	F	**.28**	.12; .42	**.25**	.17; .32	**.25**	.17; .32	.06	–.10; .22
	M	**.50**	.29; .67	**.27**	.17; .37	**.27**	.17; .37	**.34**	.10; .54
Power	F	**.42**	.28; .54	**.43**	.36; .49	**.42**	.35; .48	**.44**	.30; .56
	M	**.54**	.34; .70	**.42**	.33; .51	**.42**	.33; .51	**.34**	.10; .54

Note. Health p. = health protection, achiev. = achievement, p. growth = personal growth, F = females, M = males. Correlations printed in boldface were significant at $p \leq .001$.

[a] Correlations across informant 1 and informant 2 ratings.

[b] Females: n = 152–154, males: n = 62.

[c] Females: n = 613–614, males: n = 317–318.

[d] Correlation was significant at $p \leq .02$.

Table A3

Regression Models on the (Moderating) Effect of Gender on Self-Other Agreement

Informant	Self-rating			Gender			Self-rating × gender			Model	
rating	*b*	*SE*	*p*	*b*	*SE*	*p*	*b*	*SE*	*p*	*F*	*p*
Protective											
Health p.	**0.48**	0.05	<.001	0.08	0.06	.148	**–0.15**	0.06	.011	69.27	.004
Intimacy	**0.41**	0.04	<.001	**0.27**	0.06	<.001	–0.02	0.06	.766	81.42	<.001
Expansive											
Achiev.	**0.37**	0.05	<.001	–0.09	0.06	.121	–0.01	0.06	.836	61.31	<.001
Affiliation	**0.60**	0.04	<.001	**0.14**	0.05	.010	**–0.14**	0.05	.009	154.93	<.001
P. growth	**0.30**	0.05	<.001	0.07	0.06	.252	–0.06	0.06	.348	26.25	<.001
Power	**0.43**	0.04	<.001	**–0.21**	0.06	<.001	–0.02	0.06	.785	92.89	<.001

Note. Health p. = health protection, achiev. = achievement, p. growth = personal growth, b = unstandardized regression weight, *SE* = standard error of estimate, df = 3 in all models. Motive scales were standardized prior to the analyses. Significant model parameters ($p < .05$) are printed in boldface.

Table A4

Regression Models on the (Moderating) Effect of Relationship Duration on Self-Other Agreement

Informant rating	Self-rating			Rel. duration			Self-rating × rel. duration			Model	
	b	*SE*	*p*	*b*	*SE*	*p*	*b*	*SE*	*p*	*F*	*p*
Protective											
Health p.	**0.39**	0.03	<.001	**0.06**	0.03	.048	–0.00	0.03	.944	67.54	<.001
Intimacy	**0.39**	0.03	<.001	**–0.13**	0.03	<.001	–0.04	0.03	.120	82.69	<.001
Expansive											
Achiev.	**0.37**	0.03	<.001	–0.02	0.03	.585	–0.02	0.03	.545	60.72	<.001
Affiliation	**0.52**	0.03	<.001	–0.04	0.03	.091	–0.03	0.03	.280	149.79	<.001
P. growth	**0.27**	0.03	<.001	–0.03	0.03	.246	**–0.07**	0.03	.030	27.67	<.001
Power	**0.44**	0.03	<.001	**0.09**	0.03	.001	0.02	0.03	.557	91.70	<.001

Note. Rel. duration = relationship duration, health p. = health protection, achiev. = achievement, p. growth = personal growth, *b* = unstandardized regression weight, *SE* = standard error of estimate, *df* = 3 in all models. Relationship duration and motive scales were standardized prior to the analyses. Significant model parameters ($p < .05$) are printed in boldface.

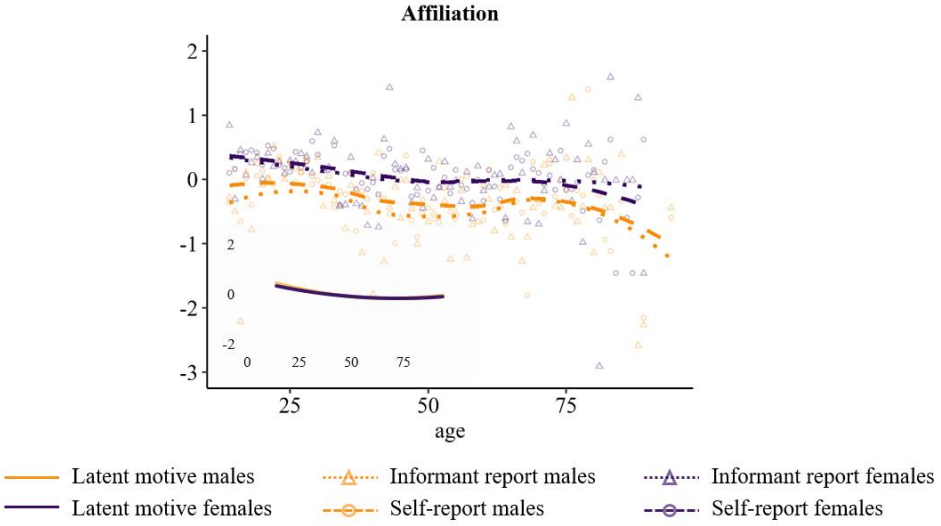

Figure A1. Alerting mean-level trends separated by gender, assessed by self-reports and averaged informant reports; and the corresponding latent mean-level trends based on self-other composite scores in affiliation including age² effects (gray background).

Appendix II

Transactions Between Self-Esteem and Conflict in Romantic Relationships — A Five-Year Longitudinal Study

Julia Richter & Christine Finn

Abstract

Self-esteem has been shown to be both predictive of and predicted by characteristics of romantic relationships. To date, most studies on the dynamics of self-esteem in romantic couples have focused on positive relationship aspects, such as satisfaction. Yet, to embrace relationship dynamics more comprehensively, it is essential to also investigate negative aspects, such as conflict. To that aim, we examined the transactional and longitudinal interplay between self-esteem and perceived relationship conflict in stable romantic couples. Our sample consisted of N = 1,093 young adult female–male relationships from the German Family Panel. Individuals' self-esteem, perceived conflict frequency, and their perception of their partners' dysfunctional conflict styles were examined annually throughout a time span of five years. Based on dyadic bivariate latent change models, we tested our assumption that self-esteem and aspects of perceived relationship conflict are negatively interrelated within individuals and between partners both within and across time. Actor effects highlighted the importance of perceived conflict frequency for subsequent self-esteem changes, beyond and above self-esteem levels. Transactions between changes in conflict frequency and in self-esteem indicated that both domains are longitudinally intertwined. Moreover, we found actor and partner effects of perceived partner conflict styles on self-esteem changes. The results imply that perceiving conflict is more important for the development of self-esteem than vice versa, thereby supporting sociometer perspectives. The lack of support for broadcasting perspectives with regard to perceived partner conflict styles indicate that negative and positive relationship dynamics might not function parallel to each other. In addition, consequences of dysfunctional conflict behavior are not a within-person experience only, but affect both partners' self-esteem.

Introduction

As romantic relationships might be the closest relationships that adults in Western societies can have [1], they might seem to be a perfect context for studying personality–relationship transactions [2]. Focusing much of its attention on the forming and succeeding of partnerships, previous research has found that personality traits, and especially self-esteem, are strong predictors of relationship characteristics [2,3,4]. Moreover, having or establishing a partnership can have positive effects on self-esteem changes [2,5,6]. These studies have established that self-esteem and positive relationship characteristics increase one another.

However, romantic relationships comprise multiple aspects [7] of positive and negative evaluative character. To understand partner dynamics more extensively, it is necessary to also investigate negative self-esteem–relationship dynamics. One negative relationship characteristic is experiencing relationship conflict [8]. "*Conflict can be defined as the perception of contention or disaccord*" [9, p. 13]. Only a few studies to date have explicitly investigated long-term transactions between self-esteem and perceived conflict in dating couples. With our study, we filled this gap, examining dyadic transactions between self-esteem and multiple indicators of perceived relationship conflict in stable romantic relationships throughout a time span of five years.

Self-Esteem and Romantic Relationships

Global self-esteem refers to an individual's overall self-evaluation of being a person of worth [10]. Being moderately stable across time [11,12], some changes in self-esteem can be explained by relationship experiences [2,13]. There are several perspectives that provide explanations on why self-esteem and romantic relationship experiences might be longitudinally intertwined with each other.

Sociometer perspectives on self-esteem. The sociometer perspective states that self-esteem is dependent on social embeddedness, acting as a "*sociometer that monitors the degree to which the individual is being included versus excluded by other people and that motivates the person to behave in ways that minimize the probability of rejection or exclusion*" [14, p. 518]. According to this theory, the function of low self-esteem is to indicate failure in maintaining an adaptive social and interpersonal inclusion as prerequisite for survival and reproduction. Besides implying that romantic relationships should be the best context to study this perspective, sociometer theory assumes that perceiving rejection (i.e., a fight, withdrawal) in a partnership should lead individuals to decrease in self-esteem.

Let us assume that we accompany the couple Jenny and James. We could imagine several scenarios that are explainable with the sociometer perspective. First, if Jenny and James perceive multiple conflict situations within their relationship, this could lead both of them to feel rejected, leading to decreases in both their self-esteem. Second, perceiving his partner Jenny showing withdrawal behavior might lead James to think that she rejects him, which decreases his self-esteem. And, third, if James constantly perceives and reports that Jenny shows dysfunctional, e.g., destructive, behavior, this might tell something about James' devaluation of Jenny. Put differently, James might not appreciate Jenny's behavior, which leads her to feel rejected, and in turn might lead her to decrease in self-esteem.

Self-broadcasting perspectives on self-esteem. The self-broadcasting perspective refers to the effect of an individual's self-evaluation on socially relevant outcomes [15]. This perspective assumes that observable behavior of individuals in a social context is directly dependent on their self-evaluations. Because individuals behave in accordance with their self-evaluations, they consequently evoke the expected (positive or negative) reactions in their social encounters [16]. The risk (or dependency) regulation model posits similar assumptions [17]. It states that internal beliefs about the partner's regard, which depend on a person's self-esteem, affect their confidence to seek dependence and connectedness [18]. The authors [17] argued that "*[r]ather than accurately mirroring the partner's view of the self, perceptions of a partner's regard may largely be a projection, reflecting self-perceived worthiness of love*" (p. 479). Put differently, low self-esteem individuals will come to negative assumptions about how their partners view them, and, in turn, their actual relationship.

For example, James, who has low self-esteem, might perceive more negative conflict behavior in his partner Jenny than a high self-esteem individual would do. Moreover, low self-esteem individuals should distance themselves from their partners (in terms of avoidance, for example) to minimize the risk of rejection, whereas high self-esteem individuals should risk and seek closeness, even after conflict situations [18]. This would eventually lead partners of low self-esteem individuals to perceive them showing avoiding behavior, which is mirroring the low self-esteem individuals' attempt for self-protection.

Support for Sociometer and Self-Broadcasting Perspectives in Romantic Relationships

In accordance with sociometer perspectives, which see the function of low self-esteem as an indicator of rejection, anxious and avoidant attachment behavior predicted lower self-evaluations [15]. Moreover, self-esteem has been found to increase after perceiving approving social reactions [19]. To conclude, changes in self-esteem can be triggered by social

experiences within relationships. Yielding support for the self-broadcasting perspective, small positive effects of self-esteem on relationship satisfaction changes have been reported [20,21]. The authors [2] addressed the dynamic transactions between self-esteem and positive relationship aspects using a dyadic perspective [22]. That is, besides assessing self-esteem and perceived positive relationship aspects within one partner, they also assessed the same constructs in the other partner, too. Applying a dyadic approach did not only allow them to examine personality–relationship transactions within individuals (*actor effects*). Rather, including both partners in one model made it possible to examine between-person dynamics of these transactions (*partner effects*). Supporting the notion that person and relationship characteristics influence each other [23], they found hints for both perspectives, both within individuals and between partners [2].

Self-esteem and perceived conflict in romantic relationships. Murray and colleagues [19] reported actor and partner effects showing that higher self-esteem predicted decreases in destructive conflict behavior, but not in conflict frequency. They did not explicitly specify how they operationalized conflict. This is, however, not a trivial question. Indeed, one study suggested that one person's self-esteem is most strongly associated to their self-perception of social indicators, but not necessarily to others' perception of these indicators [24]. Moreover, although previous research suggested that self-esteem and perceived conflict might influence each other, there hardly is evidence on the longitudinal transactions behind these influences.

The Present Study

With our study, we aim to broaden the knowledge about self-esteem changes in accordance with negative relationship dynamics. To that end, we examined self-esteem–conflict transactions in stable romantic female–male couples annually across a five-year time span. Using dyadic bivariate latent-change models [following the procedure by 2], we examined long-term associations between both partners' self-esteem and perceived conflict frequency. In addition, we considered perceived dysfunctional conflict styles in terms of the partner's destructive behavior, and withdrawal. A dyadic longitudinal study within stable partnerships is appropriate because research has shown that relationship-specific characteristics and interactions are best portrayed by the perspectives of both partners in a dyadic interplay [22]. Moreover, our approach allowed to capture self-esteem–relationship transactions in a stable social context, making it possible to examine within-person as well as between-person dynamics. We tested the hypotheses that one person's self-esteem influenced

their perception of (dysfunctional) conflict behavior and vice versa (actor effects). At the same time, we could investigate if an individual's self-esteem affected their partner's perception of conflict behavior or vice versa (partner effects).

Hypotheses

We examined the assumption that self-esteem and perceived relationship conflict are negatively intertwined within individuals and between partners both within and across time. More specifically, in line with the sociometer and self-broadcasting perspectives, we had the following hypotheses. First, we expected that higher self-esteem would be associated cross-sectionally with lower perceived conflict frequency. We expected the same for lower perceived destructive behavior and withdrawal in one's partner (*correlations*). Second, we expected that an increase in self-esteem should be associated with a decrease in perceived conflict frequency and behavior within the same time span, and vice versa (*correlated changes*). For example, if James' self-esteem increased between T1–T2, his perception of his partner Jenny's destructive and withdrawal behavior should decrease during the same time span, too. As this scenario only refers to processes within one person, this would be called an actor effect. A partner effect, in contrast, would occur if these processes spilled over between partners, for example, if an increase in Jenny's self-esteem between T1–T2 is associated to James' notion that Jenny behaves less destructive and withdrawing in a conflict situation at T2 than at T1.

If Jenny perceives high conflict frequency with her partner James, this might lead her to feel disapproved, and, thus, insecure. The same should be true if Jenny constantly perceives James' conflict behavior as destructive or withdrawing. Put differently, we expected that higher levels of perceived (dysfunctional) conflict behavior in one's partner predicts decreases in an individual's self-esteem (*actor level-change effect*). We assumed that this process might spill over between partners (*partner level-change effect*). Put differently, if James constantly perceives and reports that Jenny shows (dysfunctional) conflict behavior, this might be a hint on James potential devaluation of her, and Jenny's self-esteem might decrease in reaction to this [19].

In accordance with the risk rejection hypothesis, we expected self-esteem to be negatively associated with subsequent changes in perceived (dysfunctional) conflict behavior. For example, James' low self-esteem might lead him to increasingly perceive dysfunctional conflict behavior in Jenny over time (*actor level-change effect*). Moreover, we again tested if this assumption held between partners (*partner level-change effect*). In addition to level-

change effects, we expected that an increase in perceived conflict behavior might lead to a later decrease in self-esteem and vice versa (*change-change effects*) [25]. Again, we predicted these effects both within individuals (*actor change-change effects*) and between partners (*partner change-change effects*).

Materials and Methods

For the present study, we used data from the German Family Panel (pairfam) [26]. Pairfam is a representative longitudinal study, which incorporates a multi-actor design with yearly assessments [for a detailed description, see 27], allowing for a dyadic perspective. Data collection of pairfam started in 2008, and is in consistence with the ethical standards for the treatment of human subjects (see https://www.pairfam.de/fileadmin/user_upload/redakteur/publis/Dokumentation/Formulare/letter_-19016KH_-_pairfam.pdf). At the time we started our study, data of nine waves (waves 1–9) were available.

Participants and Procedure

At wave 1 (T0), N = 12,402 anchors participated. Of those, n = 6,373 were females, and n = 6,027 were males. For the purpose of our gender-controlled dyadic design, we excluded females and males with same-sex partnerships, and persons who did not reliably indicate their gender. Moreover, we excluded anchors of the youngest cohort (birth years 1991–1993) because research has shown that adolescent romantic relationships differ qualitatively from romantic relationships later in adulthood [28], so the age groups may not be readily comparable.

Partners' self-esteem was only measured from wave 7 (T6) on if the relationship was new, which made the dyadic perspective for stable relationships impossible after wave 6 (T5). To make our sample more comparable in terms of longitudinal dynamics, we excluded pairs that had not been together at any of the assessment waves T0–T5. That is, we excluded participants that were single and/or changed their partners within this time span, and couples that were together both at T0 and T5, but indicated a separation phase in between (i.e., at one of the measurement occasions T1–T4). In wave 1 (T0), anchors' self-esteem was measured by a personal interview, but since wave 2 (T1), it was assessed via a computer-assisted self-administered interview. As the measurement method had been shown to affect the actual self-esteem measures [2], we decided to exclude wave 1 (T0) from our analysis [compare 3]. We thus conducted all analyses across T1–T5. Little's MCAR test with all manifest items

indicated that data might not be missing at random [$\chi^2(17{,}893) = 21{,}719.21$, $p < .001$]. As the χ^2-to-*df* ratio of 1.20 revealed, however, this could be due to the large sample size [compare 2].

Of course, our sample was biased in the way that only stable couples' interactions were followed. At T1, our dataset consisted of $N = 1{,}093$ female–male couples (i.e., 1,093 females and 1,093 males, respectively) with females reporting a mean age of 31.72 years ($SD = 5.31$, age range: 18–52 years), and males a mean age of 34.55 years ($SD = 6.03$, age range: 20–68 years). For most relationships, anchors and their partners (here, we report on the information given by the anchors) indicated to be married/in a civil union ($n = 796$, 72.8%). Mean relationship duration at T1 was eleven years (i.e., $M = 127.77$ months, $SD = 68.32$ months, duration range: 12–403 months, 1 information missing). The mean number of children was 1.37, ranging from 0–7 children ($SD = 1.12$). At T1, $n = 801$ (36.6%) participants had absolved the general higher education qualification ("*Allgemeine Hochschulreife*").

Measures

Self-esteem. Self-esteem was assessed via a computer-assisted self-administered interview at all measurement occasions (T1–T5). Three items from the Rosenberg Self-Esteem Scale [29] (e.g., "*I like myself just the way I am*") were answered on a five-point Likert scale ranging from 1 (*not at all*) to 5 (*absolutely*). One item was reversed before the analysis. Coefficient ω and its 95% confidence intervals of the self-esteem scale for females were .78 [.75; .81], .78 [.75; .80], .78 [0.76; 0.81], .74 [.70; .77], and .79 [.76; .81] at T1–T5, respectively. Coefficient ω estimates for males were .74 [.71; .77], .74 [.70; .77], .77 [.74; .79], .71 [.66; .75], and .79 [.76; .82] at T1–T5, respectively.

Perceived conflict frequency in romantic relationships. Relationship conflict was measured with multiple indicators. First, we used two items of an adapted Conflict Scale from the Network of Relationships Inventory [8] to measure perceived *conflict frequency* within the romantic relationship from both partners' perspectives from 1 (*never*) to 5 (*always*) in all waves. An example of the English item translation would be, "*How often do you and [name of current partner] disagree and quarrel?*" Coefficient ω and its 95% confidence intervals of females' perceived conflict frequency were .78 [.74; .81], .81 [.78; .83], .79 [.76; .81], .82 [.78; .84], and .81 [.78; .84] at T1–T5, respectively. Coefficient ω estimates of males' perceived conflict frequency were .76 [.72; .79], .79 [.75; .82], .78 [.75; .81], .80 [.77; .83], and .79 [.76; .82] at T1–T5, respectively.

Perceived conflict styles in romantic relationships. In contrast to perceived conflict frequency, we assessed perceived conflict styles of the romantic partner within conflict situations that we hypothesized to be associated to rejection perception. More specifically, we assessed individuals' perceptions *of their partners'* behavior within conflict situations. We recoded perceived constructive behavior from an adapted Constructive Behavior Scale [30] into *perceived destructive behavior* to address the negativity and unwillingness to compromise in partner conflicts. *Perceived withdrawal behavior* of the partner was assessed via the Withdrawal Scale of the Conflict Resolution Inventory [31]. Both scales were assessed on 5-point scales from 1 (*almost never or never*) to 5 (*very frequently*). The introductory question to individuals' perceptions of their partners' conflict styles was, "*How often did your partner engage in any of these behaviors?*" An example item for perceived constructive (destructive) behavior in one's partner would be "*listen to and ask questions of you in order to understand better* (item recoded)." An example for the perceived tendency in one's partner to withdrawal would be "*remain silent.*"

Coefficient ω of females' perception of their male partners' destructive behavior were .72 [.68; .76], .68 [.63; .71], .76 [.72; .79], .72 [.68; .76], and .75 [.71; .78] at T1–T5, respectively. Coefficient ω and its 95% confidence intervals of males' perception of their female partners' destructive behavior were .66 [.61; .70] at T1, .68 [.63; .72] at T2 and T3, .68 [.63; .72] at T4, and .70 [.65; .74] at T5. Coefficient ω of females' perception of their male partners' withdrawal were .56 [.51; .61], .61 [.56; .65], .63 [.58; .67], .68 [.63; .72], and .66 [.62; .71]. Coefficient ω and its 95% confidence intervals of males' perception of their female partners' withdrawal were .61 [.56; .65], .63 [.58; .68], .64 [.59; .69], .68 [.63; .72], and .65 [.59; .69] at T1–T5, respectively.

Analytical Strategy

We examined transactions between both partners' self-esteem and aspects of perceived conflict in three separate adapted longitudinal actor–partner (i.e., dyadic) interdependence models [2,22] (i.e., one for perceived conflict frequency, perceived partner destructive behavior, and perceived partner withdrawal, respectively). The use of dyadic bivariate latent change models allowed us to test all of our hypotheses per conflict variable in one model. That is, we were able to assess transactions between self-esteem and aspects of perceived relationship conflict across time and between partners in terms of initial correlations, correlated changes, level-change effects and change-change effects, above and beyond the

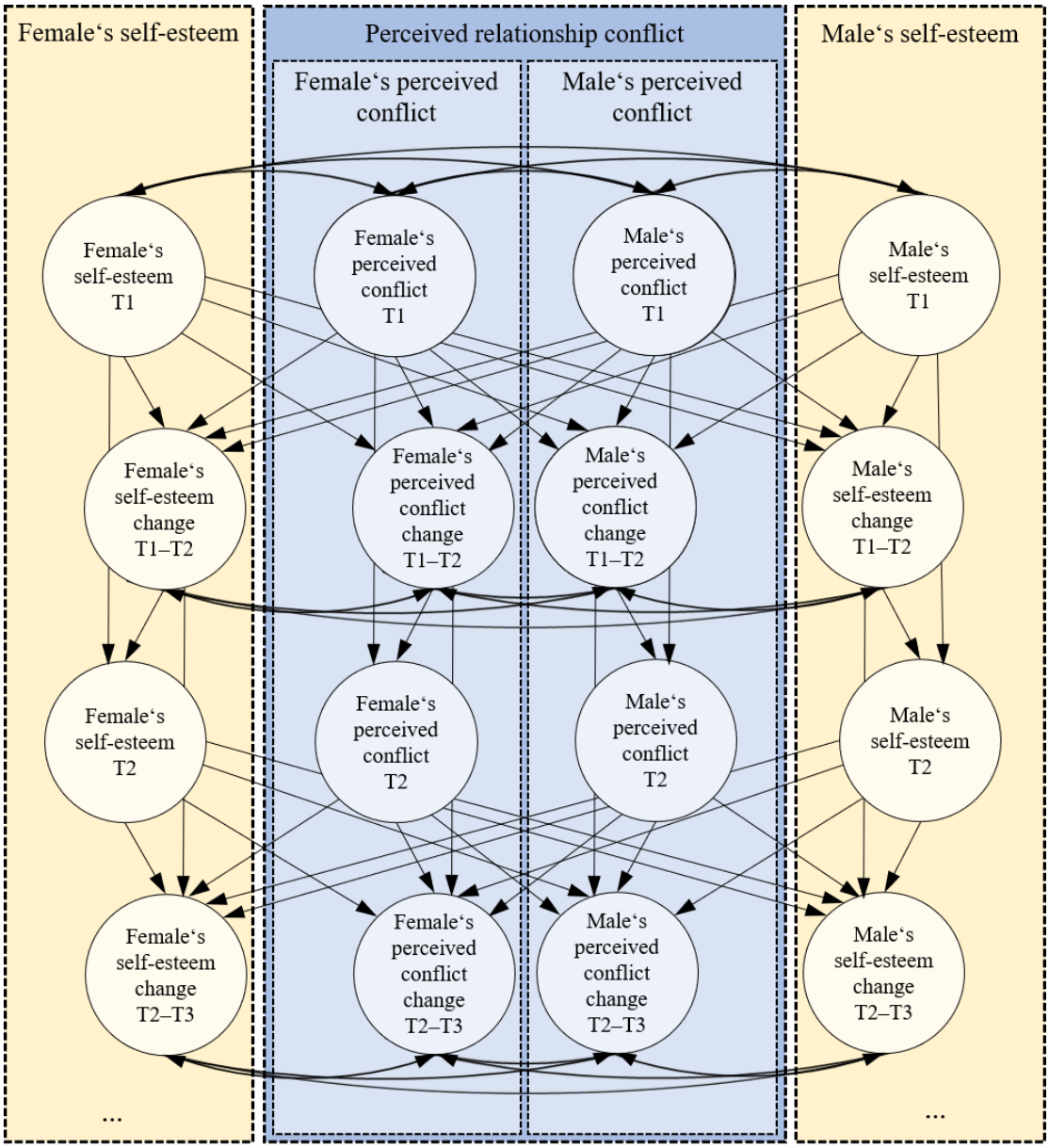

Figure 1. Dyadic extended bivariate latent change model of self-esteem and perceived relationship conflict. For parsimony, we only printed exemplary time points T*n* instead of the five measurement occasions of our model. Measurement models and paths were equated across genders and time periods. Measurement models are omitted for parsimony. Adapted from [2].

control of measurement errors, interrelations, and prior variable levels (see Figure 1). In the following, we will describe the procedure in more detail.

Latent change modeling. To get the most reliable measures, we modeled each anchor’s and partner’s self-esteem and perceived relationship conflict as latent variables, i.e., controlling for measurement error [32]. We first established univariate models to individually capture self-esteem and perceived conflict change patterns of females and males in stable relationships. To assess latent changes, variable levels were decomposed into the initial variable level, i.e., level at T1, and the latent variable changes, i.e., the differences between latent trait levels at two neighboring measurement occasions [33,34]. We established strong

measurement invariance by equating the latent variable structure, factor loadings, and intercepts for each construct across time in order to reasonably interpret the meaning of their mean-level changes [35]. Moreover, we equated change parameters across time to get the most parsimonious model. Residual variances of each item within a construct were allowed to correlate across time points to account for shared variance that was not accounted for by the latent variables [36].

Dyadic bivariate transactions. In a second step, we integrated the univariate models into the dyadic bivariate design. That is, we combined the latent change models for females' and males' self-esteem with the respective latent change models for aspects of females' and males' perceived conflict. These models allowed us to observe initial correlations between the constructs at T1 within individuals and across partners. Correlated changes, that is, the correlation between change scores within concurrent time intervals, denote the degree to which constructs develop into similar or divergent directions, again both within and across partners. Level-change effects describe how levels in one construct at T1 predict subsequent changes in another construct. Last, change-change effects describe how prior change in one construct predicts subsequent change in another [37]. Actor and partner effects in level-change and change-change effects from perceived relationship conflict on self-esteem made it possible to test the sociometer hypothesis, while effects from self-esteem on perceived relationship conflict pertain to the self-broadcasting perspective. We included age and relationship duration as covariates of all effects in our analyses. To yield the most parsimonious model, we equated paths for women and men [2,21].

Data preparation was done using R version 3.4.4 [38]. The dyadic bivariate latent change models were implemented and run in Mplus version 8 [39], where missing data was handled using FIML [40]. All syntax files can be found at the open science framework: https://osf.io/rc9qf/?view_only=a89ae9928223400986a97993a64579d6 (*This is currently a view-only link that will be completely open after acceptance*). Combining CFI, RMSEA, and SRMR leads to a good overall model fit criterion for large sample sizes. As a rule of thumb, CFI (comparative fit index) should be > .93, and both RMSEA (root mean square error of approximation) and SRMR (standardized root mean square residual) < .07, respectively [41].

Results

Univariate Change Statistics

Table 1

Descriptive statistics of the study variables

	M (*SD*)					Cohen's *d* [a] [95% CI]				Stability [b] [95% CI]			
Variable	**T1**	**T2**	**T3**	**T4**	**T5**	**T1→T2**	**T2→T3**	**T3→T4**	**T4→T5**	r_{12}	r_{23}	r_{34}	r_{45}
							Reported by females						
Self-esteem	3.81 (0.86)	3.82 (0.84)	3.85 (0.82)	3.85 (0.81)	3.85 (0.80)	.01 [–.07; .10]	.04 [–.04; .12]	.00 [–.08; .08]	.00 [–.08; .08]	.56 [.52; .60]	.59 [.55; .63]	.61 [.57; .65]	.64 [.60; .67]
Conflict frequency	2.56 (0.61)	2.58 (0.64)	2.59 (0.62)	2.57 (0.63)	2.58 (0.63)	.04 [–.04; .12]	.02 [–.07; .10]	–.04 [–.12; .04]	.02 [–.06; .10]	.66 [.62; .69]	.64 [.61; .68]	.68 [.64; .71]	.70 [.67; .73]
Destructive behavior	2.49 (0.97)	2.56 (0.91)	2.53 (0.96)	2.59 (0.92)	2.64 (0.91)	.07 [–.01; .16]	–.03 [–.12; .05]	.07 [–.02; .15]	.06 [–.03; .14]	.48 [.43; .53]	.52 [.47; .56]	.55 [.50; .59]	.55 [.51; .59]
Withdrawal	2.33 (1.02)	2.35 (1.04)	2.29 (1.02)	2.36 (1.03)	2.33 (1.01)	.02 [–.06; .11]	–.06 [–.15; .02]	.08 [–.01; .16]	–.03 [–.12; .05]	.62 [.58; .65]	.58 [.53; .62]	.61 [.57; .65]	.61 [.57; .65]
							Reported by males						
Self-esteem	4.03 (0.75)	4.05 (0.73)	4.05 (0.75)	4.03 (0.72)	4.06 (0.74)	.03 [–.06; .11]	.00 [–.08; .08]	–.03 [–.11; .06]	.05 [–.04; .13]	.50 [.45; .54]	.51 [.46; .56]	.51 [.47; .56]	.57 [.53; .61]
Conflict frequency	2.50 (0.62)	2.47 (0.63)	2.45 (0.59)	2.48 (0.62)	2.44 (0.61)	–.06 [–.14; .03]	–.04 [–.12; .05]	.06 [–.03; .14]	–.08 [–.16; .01]	.62 [.58; .66]	.63 [.59; .67]	.63 [.59; .67]	.63 [.59; .66]
Destructive behavior	2.35 (0.84)	2.45 (0.86)	2.41 (0.84)	2.52 (0.84)	2.50 (0.81)	.11 [.03; .19]	–.04 [–.13; .04]	.13 [.04; .21]	–.02 [–.11; .06]	.42 [.36; .47]	.41 [.35; .46]	.46 [.41; .51]	.48 [.43; .53]
Withdrawal	2.08 (0.91)	2.04 (0.91)	2.02 (0.88)	2.11 (0.92)	2.03 (0.88)	–.05 [–.13; .04]	–.02 [–.11; .06]	.11 [.03; .19]	–.10 [–.18; –.01]	.56 [.52; .60]	.56 [.51; .60]	.57 [.52; .61]	.60 [.55; .64]

Note. [a] for samples with repeated measures; [b] subscripts denote measurement occasions. Conflict frequency = perceived conflict frequency; destructive behavior = perceived destructive behavior tendencies in partner; withdrawal = perceived withdrawal tendencies in partner. Adapted from [2].

All correlations were significant at $p < .01$.

We did not find any substantial mean-level changes in self-esteem, perceived conflict frequency, and perceived conflict styles (see Table 1). Exceptions were males' perception of their female partners' destructive behavior, increasing across T1–T2 (d = .11, 95% CI [.03; .19]) and T3–T4 (d = .13, 95% CI [.04; .21]), and males' perception of their female partners' withdrawal, also increasing across T3–T4 (d = .11, 95% CI [.03; .19]). All variables showed moderate rank-order consistencies for both genders (see Table 1), which indicated individual variability in trait change trajectories. Initial within-construct correlations and correlated changes between partners can be found in S1 Table, within-construct partner level-change and change-change effects in S2 Table.

Transactions Between Self-Esteem and Relationship Conflict

All models indicated a good fit to the data with RMSEA = .027 [90% CI: .025; .028], CFI = .965, SRMR = .057 for conflict frequency; RMSEA = .026 [90% CI: .024; .028], CFI = .957, SRMR = .053 for partner destructive behavior; and RMSEA = .026 [90% CI: .024; .028], CFI = .960, SRMR = .058 for partner withdrawal. Initial correlations between self-esteem and aspects of perceived relationship conflict were negative within individuals and between partners in all cases (r = –.27 to –.13, p < .01, see Table 2). Changes between self-esteem and all aspects of perceived conflict were also negatively correlated both within individuals and between partners (r = –.23 to –.06, p < .01). For example, decreases in one individual's self-esteem were associated with concurrent increases in their partner's perceived conflict frequency. We have to stress, however, that all correlations were in a low range.

Level-change effects supporting self-broadcasting perspectives. We found one marginally significant negative actor effect of self-esteem level on subsequent change in perceived partner destructive behavior (b = –.05 [95% CI: –.11; .00], p = .055). That is, higher initial self-esteem predicted comparably steeper subsequent decrease in the perception of the partner's destructive behavior, above and beyond the control of initial correlations and prior levels. The other way around, lower initial self-esteem predicted comparably steeper subsequent increase in perceived partner destructive behavior. In all other cases, we did not find support for the self-broadcasting or the risk regulation perspectives in terms of level-change effects (see Table 3).

Table 2

Initial correlations and correlated changes between self-esteem and aspects of perceived relationship conflict

	Self-esteem									
	T1		**T1→T2**		**T2→T3**		**T3→T4**		**T4→T5**	
Conflict variable	***r***	**95% CI**	***r***	**95% CI**	***r***	**95% CI**	***r***	**95% CI**	***r***	**95% CI**
Within individuals (*actor effects*)										
Frequency	–.27	–.32; –.22	–.18	–.22; –.14	–.20	–.24; –.15	–.23	–.28; –.18	–.19	–.23; –.15
Destructive behavior	–.13	–.19; –.08	–.11	–.14; –.07	–.11	–.15; –.08	–.14	–.19; –.10	–.12	–.16; –.08
Withdrawal	–.23	–.29; –.17	–.17	–.22; –.13	–.18	–.23; –.13	–.20	–.25; –.15	–.18	–.22; –.13
Between partners (*partner effects*)										
Frequency	–.17	–.22; –.11	–.09	–.12; –.05	–.10	–.14; –.06	–.12	–.16; –.07	–.09	–.13; –.06
Destructive behavior	–.13	–.18; –.07	–.06	–.09; –.02	–.06	–.10; –.03	–.08	–.12; –.03	–.07	–.10; –.03
Withdrawal	–.19	–.25; –.13	–.09	–.13; –.04	–.09	–.14; –.05	–.10	–.15; –.05	–.09	–.13; –.05

Note. Frequency = perceived conflict frequency; destructive behavior = perceived destructive behavior tendencies in partner; withdrawal = perceived withdrawal tendencies in partner. Coefficients were averaged across sexes due to slight differences in sex-specific variances. Adapted from [2].

All correlations were significant at $p \leq .001$.

Table 3

Dyadic latent bivariate level-change and change-change effects

Effects of self-esteem on perceived relationship conflict

Conflict variable	Self-esteem level on conflict change [a]						Self-esteem change on conflict change [b]					
	Within individuals (*actor effects*)			Between individuals (*partner effects*)			Within individuals (*actor effects*)			Between individuals (*partner effects*)		
	b	*p*	95% CI	*b*	*p*	95% CI	*b*	*p*	95% CI	*b*	*p*	95% CI
Frequency	–.01	.625	–.05; .03	.01	.657	–.03; .05	–.08	.020	–.15; –.01	.02	.577	–.05; .09
Destructive behavior	–.05	.055	–.11; .00	–.03	.230	–.09; .02	.04	.321	–.04; .12	–.05	.275	–.13; .04
Withdrawal	.02	.467	–.03; .06	–.01	.511	–.06; .03	–.01	.865	–.07; .06	.03	.428	–.04; .10

Effects of perceived relationship conflict on self-esteem

Conflict variable	Conflict level on self-esteem change [a]						Conflict change on self-esteem change [b]					
	Within individuals (*actor effects*)			Between individuals (*partner effects*)			Within individuals (*actor effects*)			Between individuals (*partner effects*)		
	b	*p*	95% CI	*b*	*p*	95% CI	*b*	*p*	95% CI	*b*	*p*	95% CI
Frequency	–.03	.001	–.04; –.01	.00	.651	–.01; .02	.05	.003	.02; .08	–.02	.121	–.05; .01
Destructive behavior	–.01	.066	–.02; .00	–.01	.022	–.03; –.00	–.01	.180	–.04; .01	.01	.253	–.01; .03
Withdrawal	–.02	.050	–.03; .00	–.04	<.001	–.05; –.02	–.02	.422	–.05; .02	.02	.224	–.02; .06

Note. Frequency = perceived conflict frequency; destructive behavior = perceived destructive behavior tendencies in partner; withdrawal = perceived withdrawal tendencies in partner. For parsimony, unstandardized coefficients were equated across sexes and across time intervals. Adapted from [2].

[a] T1 on T1 → T2, T2 on T2 → T3, etc.

[b] T1 → T2 on T2 → T3, T2 → T3 on T3 → T4, etc.

Level-change effects supporting sociometer perspectives. We found a small within-person (actor) effect of perceived conflict frequency level on subsequent self-esteem change (b = –.03 [95% CI: –.04; –.01], p = .001), but no partner effect. Marginally significant actor effects of perceived partner destructive behavior (b = –.01 [95% CI: –.02; .00], p = .066) and withdrawal (b = –.02 [95% CI: –.03; .00], p = .050) on subsequent self-esteem changes tended to support sociometer perspectives for conflict styles. That is, the more one individual perceived destructive behavior or withdrawal in their partner, the steeper their self-esteem tended to decrease. Significant partner effects of perceived partner destructive behavior (b = –.01 [95% CI: –.03; –.00], p = .022) and withdrawal (b = –.04 [95% CI: –.05; –.02], p < .001) on subsequent changes in the partner's self-esteem indicated the following: The more an individual perceived destructive or withdrawal tendencies in their partner, the steeper their partner's self-esteem decreased consequently.

Change-change effects. On the actor level, there were two significant change-change effects: Increases in perceived conflict frequency predicted subsequent increases in self-esteem (b = .05 [95% CI: .02; .08], p = .003). To the contrary, increases in self-esteem predicted decreases in perceived conflict frequency (b = –.08 [95% CI: –.15; –.01], p = .020). This pattern indicates longitudinal transactions between self-esteem and conflict frequency. We found no change-change effects for perceived destructive behavior or withdrawal, neither within individuals nor between partners.

Discussion

The aim of this study was to deepen knowledge on negative relationships dynamics in stable romantic couples. To that end, we analyzed longitudinal transactions between self-esteem and perceived conflict from both partners' perspectives, covering a time span of five years. In general, findings indicated that self-esteem and aspects of perceived conflict affected each other within and across time. This was not only true within individuals, but also between partners, indicating a shared negative dynamic within romantic relationships. More specifically, our findings mostly indicated support for sociometer perspectives, with no support for self-broadcasting and risk regulation perspectives. This was indicated by effects of aspects of perceived conflict on self-esteem changes, partly within and between partners. In the following, we will elaborate on the reasons for these findings.

Transactions Between Self-Esteem and Perceived Conflict in Stable Couples

Findings on self-broadcasting perspectives. We found only one marginally significant actor effect in support of the self-broadcasting perspective, indicating that lower levels of self-esteem affected an increase in perceived destructive behavior. Apart from this, we did not find any support for self-broadcasting or risk regulation perspectives. This finding is in line with recent studies [2,20,21], which reported only small self-broadcasting effects for relationship satisfaction, but diverging effects for other indicators of relationship quality. When reviewing the literature and in accordance with our findings, it might be concluded that self-broadcasting effects hold in terms of positive relationship aspects [42], but not in terms of conflict experiences [19]. Possibly, negative and positive dynamics in romantic relationships do not function parallel to each other.

One explanation might be that higher self-esteem has a positive effect on the view of the self [23,43] and the environment, which results in higher ratings of positive relationship characteristics, whereas evaluating negative relationship characteristics is mainly independent from self-esteem. This pattern of results might of course also be attributable to the constructs under investigation. For example, relationship satisfaction might be a more subjective criterion than conflict frequency. To clear up this uncertainty, future research would need to test differences between positive and negative dynamics systematically.

Support for sociometer perspectives. In general, both perceived conflict frequency and conflict styles showed small, but significant influences on self-esteem changes, supporting the sociometer perspective. This finding is a strong contrast to results of an actor-based longitudinal study that did not support this perspective regarding satisfaction in romantic relationships [20]. It is, however, in line with previous findings indicating small positive effects of relationship satisfaction aspects on both partners' self-esteem changes [2].

Turning to the within-person findings, we found negative actor effects of perceived conflict frequency level on subsequent self-esteem changes. That is, the more conflict individuals perceived in their relationship, the steeper was their decrease in self-esteem thereafter. Although effects were small, there seems to be an influence of perceived conflict frequency on the within-person level [19]. Put differently, conflict in romantic relationships may be experienced as an indicator of social rejection [14]. We found small tendencies for similar effects within individuals for perceived conflict styles of the partner. That is, individuals who perceived their partners to be more destructive or withdrawing tended to show stronger decreases in self-esteem (beyond and above the control of self-esteem on this perception). This pattern is in line with the sociometer perspective because the partner's

dysfunctional conflict styles might evoke feelings of social rejection, which may in turn affect self-esteem [14,19].

Moreover, we observed partner effects of perceived conflict styles on changes in self-esteem. Referring to our example of Jenny and James, James' perception of Jenny showing high levels of destructive and withdrawing behavior during conflict situations affected stronger decreases in Jenny's self-esteem. This finding is in line with [15] who reported that lower self-evaluations were predicted by insecure attachment behavior. It might be true that social rejection perception might have effects on the person who was showing rejection behavior in the first place (here, in terms of dysfunctional conflict styles), although this might be unintuitive. For example, Jenny might notice that James perceives her behavior to be dysfunctional. She might interpret his disapproval as rejection, leading her self-esteem to decrease [19]. Of course, our study design does not allow to uncover the underlying mechanisms behind our findings. It might, however, be fruitful if future research explicitly tested this assumption. Besides, our results indicate that negative conflict behavior (for example, during a fight) is not a within-person experience only, but seems to have similar consequences for both partners (compare effect sizes for actor and partner effects).

In total, these findings confirm the notion that romantic relationships are a meaningful context for self-esteem development, which is sensitive to both positive and negative relationship aspects. Couple consultants, for example, should bear in mind that communication and interaction behavior (of one partner) find compression in both individuals' self-esteem, which in turn impacts different areas of their lives – for example, their mental and physical health [44,45].

Limitations

This study has limitations. First, we only investigated dyads in continuing young adult female–male relationships with an initial mean relationship duration of two years who stayed together for at least five years. As low self-esteem has been found to predict relationship break-up [3], different patterns of dynamic transactions seem plausible in dissolving couples. Possibly, differently from our findings, we would find effects of self-esteem on perceived conflict behavior in such a sample. To test this assumption, it might be worthwhile to examine self-esteem–conflict transactions in couples destined to break up.

Second, perceived conflict frequency was assessed annually using self-reports, while conflict styles were assessed annually with the use of perceived reports about the partner.

Although we used partners' reports to yield a more objective perspective on the individuals' actual behavior [46,47], we cannot know if perceived conflict frequency and conflict styles were objective measures. We found a marginally negative effect of self-esteem on perceived destructive behavior in the partner. However, due to our study design, we were not able to disentangle whether this effect pertains to actual (i.e., objective) destructive behavior of the partner or whether low self-esteem affected a negatively biased perception of the partner's behavior. Nevertheless, we controlled for prior levels of both constructs, which resembles a control of these confounding measures. Moreover, lower self-esteem did not result in higher ratings of the partner's withdrawal behavior. This contradicts the assumption that a negative bias in the perception of people with lower self-esteem, i.e., rejection expectation [17], leads to a biased perception of rejection (in terms of withdrawal) behavior.

Third, we only investigated young to midlife adults' self-esteem–conflict transactions. It might be reasonable that younger or older couples differ regarding their relationship dynamics. Fourth, our study design followed one-year time intervals, and did not allow to uncover mechanisms behind negative relationship dynamics in an appropriate way. The TESSERA framework [48] suggests to decompose complex social situations into smaller units. This option can be applied to relationship situations that trigger and follow conflict behavior. This way, it might be possible to uncover mechanisms behind conflict situations that lead to self-esteem changes. Although the framework explicitly includes reactions from others, it is best applied to *one* individual. Integrating this framework into the PERSOC framework [49] might help to disentangle micro-processes underlying the sociometer (and possible self-broadcasting) effects that we found in our study from *both* partners' perspectives in dyads.

Finally, other authors have established a p level of <.01 [25,50,51] to account for the risk of inflating Type-I errors in multivariate analyses [52,53]. However, such error controlling procedures can lead a model to be overly conservative [53]. As our model was very complex and we wanted to ensure appropriate per-parameter power, we thus considered effects with a p level of <.05 as statistically significant. Although some effects were significant at $p < .01$, all effects were very small. That means that although true transactions between self-esteem and perceived conflict behavior are probable, they might still not be important in everyday life.

Conclusions

Our study aimed at investigating transactions between self-esteem and perceived negative relationship aspects in stable romantic couples from the perspectives of both partners. We found strong support for sociometer perspectives, although effect sizes were small. In contrast to studies focusing on relationship satisfaction, we did not find support for the self-broadcasting and risk regulation perspectives. We thus conclude that dynamics between self-esteem and perceived relationship conflict might differ from transactions with relationship satisfaction in a way that positive and negative dynamics may be not parallel to each other. As indicated by subsequent self-esteem decreases, perceived conflict frequency and perceived partner conflict styles might function as indicators of rejection. Importantly, we showed that one partner's dysfunctional conflict behavior impacts both partners' self-esteem. Our findings imply that both positive and negative relationship characteristics should be considered in studies investigating self-esteem-relationship transaction in order to consolidate findings and to derive potential mechanisms underlying these effects.

References

1. Neyer FJ, Wrzus C, Wagner J, Lang FR. Principles of relationship differentiation. Eur Psychol. 2011;16:267–77. doi: 10.1027/1016-9040/a000055
2. Mund M, Finn C, Hagemeyer B, Zimmermann J, Neyer FJ. The dynamics of self-esteem in partner relationships. Eur J Pers. 2015;29:235–49. doi: 10.1002/per.1984
3. Luciano EC, Orth U. Transitions in romantic relationships and development of self-esteem. J Pers Soc Psychol. 2017;112:307–28. doi: pspp0000109
4. Neyer FJ, Asendorpf JB. Personality–relationship transaction in young adulthood. J Pers Soc Psychol. 2001;81:1190–204. doi: 10.1037//0022-3514.81.6.1190
5. Lehnart J, Neyer FJ, Eccles J. Long-term effects of social investment: The case of partnering in young adulthood. J Pers. 2010;78:639–70. doi: j.1467-6494.2010.00629.x
6. Wagner J, Lüdtke O, Jonkmann K, Trautwein U. Cherish yourself: Longitudinal patterns and conditions of self-esteem change in the transition to young adulthood. J Pers Soc Psychol. 2013;104:148–63. doi: 10.1037/a0029680
7. Fletcher GJO, Simpson JA, Thomas G. The measurement of perceived relationship quality components: A confirmatory factor analytic approach. Pers Soc Psychol Bull. 2000;26:340–54. doi: 10.1177/0146167200265007
8. Furman W, Buhrmester D. Children's perceptions of the personal relationships in their social networks. Dev Psychol. 1985;21:1016–24. doi: 0012-1649/85/500.75
9. Peterson Tilden V, Galyen RD. Cost and conflict: The darker side of social support. West J Nurs Res. 1987;9:9–18. doi: 10.1177/019394598700900103
10. Donnellan MB, Trzesniewski KH, Robins RW. Self-esteem: Enduring issues and controversies. In: Chamorro-Premuzic T, von Stumm S, Furnham A, editors. The Wiley-Blackwell handbook of individual differences. Wiley-Blackwell; 2011. p. 718–46.
11. Kuster F, Orth U. The long-term stability of self-esteem: Its time-dependent decay and nonzero asymptote. Pers Soc Psychol Bull. 2013;39:677–90. doi: 10.1177/0146167213480189
12. Trzesniewski KH, Donnellan MB, Robins RW. Stability of self-esteem across the life span. J Pers Soc Psychol. 2003;84:205–20. doi: 10.1037/0022-3514.84.1.205
13. Orth U, Erol RY, Ledermann T, Grob A. Codevelopment of well-being and self-esteem in romantic partners: Disentangling the effects of mutual influence and shared environment. Dev Psychol. 2018;54:151–66. doi: 10.1037/dev0000400
14. Leary MR, Tambor ES, Terdal SK, Downs, DL. Self-esteem as an interpersonal monitor: The sociometer hypothesis. *J Pers Soc Psychol.* 1995;*68*:518–30. doi: 10.1037/0022-3514.68.3.518
15. Srivastava S, Beer JS. How self-evaluations relate to being liked by others: Integrating sociometer and attachment perspectives. J Pers Soc Psychol. 2005;89:966–77. doi: 10.1037/0022-3514.89.6.966

16. Swann WB Jr. The trouble with change: Self-verification and allegiance to the self. Psychol Sci. 1997;8:177–80. doi: 10.1111/j.1467-9280.1997.tb00407.x

17. Murray SL, Holmes JG, Griffin DW. Self-esteem and the quest for felt security: How perceived regard regulates attachment processes. J Pers Soc Psychol. 2000;78:478–98. doi: 10.1037//0022-3514.78.3.47

18. Murray SL, Holmes JG, Collins NL. Optimizing assurance: The risk regulation system in relationships. Psychol Bull. 2006;132:641–66. doi: 10.1037/0033-2909.132.5.641

19. Murray SL, Holmes JG, Griffin DW. The self-fulfilling nature of positive illusions in romantic relationships: Love is not blind, but prescient. J Pers Soc Psychol. 1996;71:1155–80. doi: 10.1037/0022-3514.71.6.1155

20. Orth U, Robins RW, Widaman KF. Life-span development of self-esteem and its effects on important life outcomes. J Pers Soc Psychol. 2012;102:1271–88. doi: 10.1037/a0025558

21. Erol RY, Orth U. Development of self-esteem and relationship satisfaction in couples: Two longitudinal studies. Dev Psychol. 2014;50:2291–303. doi: 10.1037/a0037370

22. Kenny DA, Kashy DA, Cook WL. Dyadic data analysis. New York, NY: Guilford Press; 2006. (Methodology in the social sciences (Kenny DA, series editor)).

23. Magnusson D. Personality development from an interactional perspective. In: Pervin LA, editor. Handbook of personality: Theory and research. New York, NY: The Guilford Press; 1990. p. 193–222.

24. Reitz AK, Motti-Stefanidi F, Asendorpf JB. Me, us, and them: Testing sociometer theory in a socially diverse real-life context. J Pers Soc Psychol. 2016;110:908–20. doi: 10.1037/pspp0000073

25. Mund M, Neyer FJ. Treating personality-relationship transactions with respect: Narrow facets, advanced models, and extended time frames. J Pers Soc Psychol. 2014;107:352–68. doi: 10.1037/a0036719

26. Brüderl J, Drobnič S, Hank K, Huinink J, Nauck B, Neyer FJ, et al.The German Family Panel (pairfam). ZA5678 Data file Version 9.1.0 [dataset]. GESIS Data Archive, Cologne; 2018. doi: 10.4232/pairfam.5678.9.1.0

27. Huinink J, Brüderl J, Nauck B, Walper S, Castiglioni L, Feldhaus M. Panel Analysis of Intimate Relationships and Family Dynamics (pairfam): Conceptual framework and design. Journal of Family Research. 2011;23:77–101. Available from: https://nbn-resolving.org/urn:nbn:de:0168-ssoar-376463

28. Collins WA, Welsh DP, Furman W. Adolescent romantic relationships. Annu Rev Psychol. 2009;60:631–52. doi: 10.1146/annurev.psych.60.110707.163459

29. Rosenberg M. Society and the adolescent self-image. Princeton: Princeton University Press; 1965.

30. Bodenmann, G. FDCT-N. Fragebogen zur Erfassung des dyadischen Copings als generelle Tendenz [FDCT-N. Questionnaire to assess dyadic coping as a general tendency]. Fribourg: Institut für Familienforschung und –beratung; 2000. German.

31. Kurdeck LA. Conflict resolution styles in gay, lesbian, heterosexual nonparent, and heterosexual parent couples. J Marriage Fam. 1994;56:705–22. doi: 10.2307/352880

32. Steyer R, Mayer A, Geiser C, Cole DA. A theory of states and traits—Revised. Annu Rev Clin Psychol. 2015;11:71–98. doi: 10.1146/annurev-clinpsy-032813-153719

33. Steyer R, Eid M, Schwenkmezger P. Modeling true intraindividual change: True change as a latent variable. Methods of Psychological Research. 1997;2:21–33. Available from: https://www.dgps.de/fachgruppen/methoden/mpr-online/issue2/art2/steyer.pdf

34. Steyer R, Partchev I, Shanahan MJ. Modeling true intradindividual change in structural equation models: The case of poverty and children's psychosocial adjustment. In: Little TD, Schnabel KU, Baumert J, editors. Modeling longitudinal and multilevel data: Practical issues, applied approaches, and specific examples. Mahwah, N.J.: Lawrence Erlbaum Associates; 2000. p. 109–26.

35. Marcusson-Clavertz D, Kjell ONE. Psychometric properties of the Spontaneous and Deliberate Mind Wandering Scales. Eur J Psychol Assess. 2019;35:878–90. doi: 10.1027/1015-5759/a000470

36. Marsh HW, Hau K-T. Assessing goodness of fit: Is parsimony always desirable? J Exp Educ. 1996;64:364–90. doi: 10.1080/00220973.1996.10806604

37. Grimm KJ, An Y, McArdle JJ, Zonderman AB, Resnick SM. Recent changes leading to subsequent changes: Extensions of multivariate latent difference score models. Struct Equ Modeling. 2012;19:268–92. doi: 10.1080/10705511.2012.659627

38. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria; 2017. Available from: https://www.R-project.org/

39. Muthén LK, Muthén BO. Mplus user's guide. 7th ed. Los Angeles, CA: Muthén & Muthén; 1998–2012. Available from: https://www.statmodel.com/download/usersguide/Mplus%20user%20guide%20Ver_7_r3_web.pdf

40. Enders CK. Analyzing longitudinal data with missing values. Rehabil Psychol. 2011;56:267–88.

41. Bagozzi RP, Yi Y. Specification, evaluation, and interpretation of structural equation models. J Acad Mark Sci. 2012;40:8–34. doi: 10.1007/s11747-011-0278-x

42. Harris MA, Orth U. The link between self-esteem and social relationships: A meta-analysis of longitudinal studies. J Pers Soc Psychol. 2019;forthcoming. doi: 10.1037/pspp0000265

43. Brockner J, Lloyd K. Self-esteem and likability: Separating fact from fantasy. J Res Pers. 1986;20:496–508. doi: 10.1016/0092-6566(86)90128-5

44. Rieger S, Göllner R, Trautwein U, Roberts BW. Low self-esteem prospectively predicts depression in the transition to young adulthood: A replication of Orth, Robins, and Roberts (2008). J Pers Soc Psychol. 2016;110:e16–22. doi: 10.1037/pspp0000037

45. Trzesniewski KH, Donnellan MB, Moffitt TE, Robins RW, Poulton R, Caspi A. Low self-esteem during adolescence predicts poor health, criminal behavior, and limited economic prospects during adulthood. Dev Psychol. 2006;42:381–90. doi: 10.1037/0012-1649.42.2.381

46. Brown JD. Evaluations of self and others: Self-enhancement biases in social judgments. Soc Cogn. 1986;4:353–76. doi: 10.1521/soco.1986.4.4.353

47. Kim H, Di Domenico SI, Connelly BS. Self–other agreement in personality reports: A meta-analytic comparison of self- and informant-report means. Psychol Sci. 2019;30:129–38. doi: 10.1177/0956797618810000

48. Wrzus C, Roberts BW. Processes of personality development in adulthood: The TESSERA framework. Pers Soc Psychol Rev. 2017;21:253–77. doi: 10.1177/1088868316652279

49. Back MD, Baumert A, Denissen JJA, Hartung F-M, Penke L, Schmukle SC, et al. PERSOC: A unified framework for understanding the dynamic interplay of personality and social relationships. Eur J Pers. 2011;25:90–107. doi: 10.1002/per.811

50. Roberts BW, Caspi A, Moffitt TE. Work experiences and personality development in young adulthood. J Pers Soc Psychol. 2003;84:582–93. doi: 10.1037/0022-3514.84.3.582

51. Sutin AR, Costa PT Jr, Wethington E, Eaton W. Turning points and lessons learned: Stressful life events and personality trait development across middle adulthood. Psychol Aging. 2010;25:524–33. doi: 10.1037/a0018751

52. Cribbie RA. Evaluating the importance of individual parameters in structural equation modeling: The need for type I error control. Pers Individ Dif. 2000;29:567–77. doi: 10.1016/S0191-8869(99)00219-6

53. Cribbie RA. Multiplicity control in structural equation modeling. Struct Equ Modeling. 2007;14:98–112. doi: 10.1207/s15328007sem1401_5

Supporting Information

Initial Correlations and Correlated Changes Within Self-Esteem and Relationship Conflict

S1 Table

Initial correlations and correlated changes between individuals (partner effects) within self-esteem and aspects of relationship conflict

	T1			T1→T2			T2→T3			T3→T4			T4→T5		
Variable	*r*	*p*	95% CI	*r*	*p*	95% CI	*r*	*p*	95% CI	*r*	*p*	95% CI	*r*	*p*	95% CI
Within the model of perceived conflict frequency															
Self-esteem	.13	<.001	.06; .20	.05	.032	.00; .09	.05	.033	.00; .10	.07	.032	.01; .14	.06	.032	.01; .11
Conflict frequency	.62	<.001	.56; .67	.28	<.001	.21; .35	.32	<.001	.24; .40	.32	<.001	.24; .40	.28	<.001	.21; .34
Within the model of perceived destructive behavior in partner															
Self-esteem	.13	<.001	.06; .20	.04	.055	–.00; .09	.05	.056	–.00; .10	.06	.055	–.00; .13	.05	.055	–.00; .10
Destructive behavior	.22	<.001	.14; .30	.08	.011	.02; .14	.08	.010	.02; .15	.10	.010	.02; .17	.09	.010	.02; .15
Within the model of perceived withdrawal in partner															
Self-esteem	.13	<.001	.06; .20	.04	.093	–.01; .08	.04	.094	–.01; .09	.06	.093	–.01; .12	.04	.093	–.01; .09
Withdrawal	.06	.198	–.03; .16	.13	.006	.04; .22	.13	.005	.04; .21	.11	.006	.03; .19	.12	.006	.03; .20

Note. Conflict frequency = perceived conflict frequency; destructive behavior = perceived destructive behavior tendencies in partner; withdrawal = perceived withdrawal tendencies in partner. Coefficients were averaged across sexes due to slight differences in sex-specific variances. Adapted from [2].

Initial correlations between partners regarding self-esteem were low ($r = .13$, $p < .001$), and so were (marginally) correlated changes (see S1 Table). Correlations of perceived conflict frequency were high at T1 ($r = .62$, $p < .001$), and correlated changes between partners indicated a mutual development in perceived conflict frequency ($r = .28$ to $.32$, $p < .001$). Although perceived destructive behavior in partners correlated at T1, changes in partner's perceived destructive behavior were independent of each other, indicating substantial between-person differences regarding changes in (perceptions of) destructive tendencies within a relationship. For withdrawal, we found the opposite pattern: No initial correlation, but significant (yet, small) positive correlations between partners' subsequent changes in this behavior ($r = .11$ to $.13$, $p \leq .006$), possibly indicating mutual reinforcement of (perceived) withdrawal tendencies. Please note that effects did not substantially differ across time intervals.

Level-Change Effects and Change-Change Effects Within Constructs

S2 Table

Level-change and change-change effects within self-esteem and perceived relationship conflict

Model of	Effects within self-esteem											
	Self-esteem level on self-esteem change [a]						Self-esteem change on self-esteem change [b]					
	Within individuals (*actor effects*)			Between individuals (*partner effects*)			Within individuals (*actor effects*)			Between individuals (*partner effects*)		
	b	*p*	95% CI	*b*	*p*	95% CI	*b*	*p*	95% CI	*b*	*p*	95% CI
Frequency	–.32	<.001	–.36; –.30	.03	.023	.00; .06	–.25	<.001	–.29; –.22	–.01	.540	–.04; .02
Destructive behavior	–.32	<.001	–.35; –.29	.03	.049	00; .05	–.27	<.001	–.30; –.23	–.01	.606	–.04; .02
Withdrawal	–.34	<.001	–.36; –.31	.02	.271	–.01; .04	–.27	<.001	–.30; –.23	–.01	.674	–.04; .03
	Effects within perceived relationship conflict											
	Conflict level on conflict change [a]						Conflict change on conflict change [b]					
	Within individuals (*actor effects*)			Between individuals (*partner effects*)			Within individuals (*actor effects*)			Between individuals (*partner effects*)		
	b	*p*	95% CI	*b*	*p*	95% CI	*b*	*p*	95% CI	*b*	*p*	95% CI
Frequency	–.18	<.001	–.21; –.16	.10	<.001	.07; .12	–.38	<.001	–.44; –.33	–.03	.290	–.08; .02
Destructive behavior	–.22	<.001	–.25; –.20	.05	.001	.03; .08	–.41	<.001	–.46; –.36	–.03	.256	–.07; .02
Withdrawal	–.13	<.001	–.16; –.11	.02	.112	–.01; .05	–.38	<.001	–.45; –.31	.05	.150	–.02; .11

Note. Frequency = perceived conflict frequency; destructive behavior = perceived destructive behavior tendencies in partner; withdrawal = perceived withdrawal tendencies in partner. For parsimony, unstandardized coefficients were equated across sexes and across time intervals. Adapted from [2].

[a] T1 on T1 → T2, T2 on T2 → T3, etc.

[b] T1 → T2 on T2 → T3, T2 → T3 on T3 → T4, etc.

We found small positive effects of an individual's self-esteem on their partner's subsequent self-esteem changes within the model of perceived conflict frequency and partner's destructive behavior ($b = .03$, $p \leq .049$), but not in the model of perceived partner's withdrawal (compare S2 Table), pointing to very small partner dynamics in self-esteem. A similar pattern emerged for perceived relationship conflict and destructive behavior. Perceived conflict frequency and perceived partner's destructive behavior reported by partner A both predicted subsequent changes in

partner B's perceived conflict frequency and in partner B's perception of partner A's destructive behavior, respectively (b = .05 to .10, $p \leq .001$). That is, the relatively more conflict one partner perceived at one time point, the more the other partner's perceived conflict frequency increased in the subsequent time interval (besides positive correlated changes, see S1 Table). Moreover, the relatively more destructive behavior one partner reported about the other partner at one time point, the more the other person increased in their perception of their partner's destructive behavior the subsequent time interval. Please note, however, that there were no substantial correlated changes between partners' perceived destructive behavior (i.e., effects were differing from zero, yet very small, see S1 Table). We did not find partner level-change effects regarding withdrawal. However, S1 Table indicates that withdrawal was mutually reinforced between partners – however, more within the same time interval, and not so much across time points. No longitudinal change-change effects were found between partners. Please note that we do not interpret actor effects within constructs.

Appendix III

Do Sojourn Effects on Personality Trait Changes Last? A Five-Year Longitudinal Study

Julia Richter, Julia Zimmermann, Franz J. Neyer, & Christian Kandler

Abstract

This study examined sojourners' long-term personality trait changes over five years, extending previous research on immediate sojourn effects. A sample of German students (N = 1,095) was surveyed thrice (T1-T3) over the course of an academic year. Sojourners (n = 498) lived abroad shortly after T1 for one or two semesters, stayers (n = 597) remained in their home country. Five years after T1, we surveyed the same participants (n = 441, 40.3%) again (T4). Beyond substantial selection effects, latent neighbor-change models revealed that small differences between sojourners' and stayers' openness, agreeableness, and neuroticism changes occurred early after sojourn-induced contextual change. Model estimates suggested sustained sojourn effects on openness and neuroticism changes thereafter, and a reversed effect on agreeableness change after return. Due to reduced power and low accuracy at T4, these estimates were not statistically significant. Based on model comparison analyses, however, we could rule out reversed effects for openness and accentuated effects for agreeableness and neuroticism as least likely. Moreover, separating short-term and long-term sojourners revealed no substantial differences, but recurring sojourn experiences tended to play a role in sustaining differences. We discuss implications for future studies on patterns of sojourn effects on personality trait changes.

Key words: Big Five personality trait changes; sojourn experience; student mobility; lasting personality trait changes; long-term sojourn effects

Do Sojourn Effects on Personality Trait Changes Last? A Five-Year Longitudinal Study

Student sojourns are one of the most frequent forms of international mobility within Europe (King & Ruiz-Gelices, 2003; Rivza & Teichler, 2007; Rodríguez González, Bustillo Mesanza, & Mariel, 2011). Recent research has revealed some small, but significant and consistent short-term and medium-term effects of sojourns on changes in students' Big Five personality traits (Greischel, Noack, & Neyer, 2016; Niehoff, Petersdotter, & Freund, 2017; Zimmermann & Neyer, 2013). It is, however, an open question whether and to what extent such changes are long-lasting. The present study attempts to provide answers, following students' personality development for more than four years after an international sojourn.

Although personality traits are defined as relatively stable individual patterns of thinking, feeling, striving, and acting across time, situations, and contexts (Kandler, Zimmermann, & McAdams, 2014), they are susceptible to change (Bleidorn et al., 2020). Trait changes are expected to occur through enduring changes in affective, cognitive, and behavioral patterns (Roberts, Hill, & Davis, 2017; Wrzus & Roberts, 2017). Systematic normative personality maturation has consistently been found in terms of average increases in conscientiousness, agreeableness, and emotional stability in young and middle adulthood (e.g., Bleidorn, Kandler, Riemann, Spinath, & Angleitner, 2009; Roberts, Walton, & Viechtbauer, 2006). This pattern can partly be explained by biological maturation as (the) one normative driving force in young adults' personality development (McCrae et al., 1999, 2000; Mõttus, 2017). Moreover, there is evidence that life events can shape individual differences in personality trait changes (Neyer & Lehnart, 2007; Neyer, Mund, Zimmermann, & Wrzus, 2014), although effects are often small and inconsistent across both traits and studies (Bleidorn, Hopwood, & Lucas, 2018; Specht, Egloff, & Schmukle, 2011). To date, research has not yet provided conclusive answers regarding the robustness and long-term sustainability of trait changes in response to life events (Bleidorn et al., 2020), and, in particular, student sojourns. It is thus an open question how sojourners' personality traits develop in comparison to non-sojourners' traits several years after their participation in international student mobility.

Long-term patterns of personality trait changes and their underlying mechanisms can only be observed when people are surveyed repeatedly using multiple assessment points over several years (Luhmann, Orth, Specht, Kandler, & Lucas, 2014). Using data from the longitudinal study "PEDES – Personality Development of Sojourners" (Zimmermann

& Neyer, 2013), we assessed sojourners' Big Five personality traits at four measurement occasions over five years, beginning before the start of their sojourn, and collected from 2009 to 2015.

Student Sojourns and Personality Trait Changes in Young Adulthood

In the present research, we focused on temporary international mobility experiences of students who were enrolled at German higher education institutions and moved abroad for a limited period of time, e.g., to complete some of their degree courses at a foreign university. Although this kind of international mobility has become generally important in the lives of students in industrialized societies (King & Ruiz-Gelices, 2003; Krings, Bangerter, Gomez, & Grob, 2008; Rivza & Teichler, 2007), individuals differ in terms of their international mobility engagement (e.g., Salisbury, Umbach, Paulsen, & Pascarella, 2009). The active seeking of individual environments according to one's individual characteristics – as it is the case with student sojourns – is a well-known phenomenon, mostly referred to as *selection* (e.g., Denissen, Ulferts, Lüdtke, Muck, & Gerstorf, 2014; George, Helson, & John, 2011; Lüdtke, Roberts, Trautwein, & Nagy, 2011; Specht et al., 2014). Indeed, individual personality characteristics can act to increase the probability of perceiving or experiencing certain kinds of life events (Headey & Wearing, 1989; Kandler, Bleidorn, Riemann, Angleitner, & Spinath, 2012). In addition, these individual characteristics themselves are susceptible to changes in response to life experiences (Denissen et al., 2014; Kandler & Ostendorf, 2016; Luhmann, Hofmann, Eid, & Lucas, 2012; Lüdtke et al., 2011; Neyer & Lehnart, 2007; Specht et al., 2014), even within a short period of time (Roberts, Luo et al., 2017). This effect is known as *socialization*.

Zimmermann and Neyer (2013) were among the first who studied socialization effects of an international sojourn experience on personality trait changes in German university students. They reported effects of student sojourns of at least 20 weeks, as evidenced by differences between sojourners and control students regarding openness, agreeableness, and emotional stability changes. Greischel et al. (2016) found comparable socialization effects of a high-school sojourn in a German sample of adolescents. Besides, Niehoff et al. (2017) reported positive effects of a college sojourn on extraversion, agreeableness, and emotional stability changes. To conclude, the overall pattern of sojourn effects on personality trait changes suggests that international sojourns contain life experiences that promote personality

maturation in young adulthood (see also Bleidorn et al., 2013; Roberts, Wood, & Smith, 2005).

Long-Term Personality Trait Changes of Sojourners

All previous studies reported personality trait changes in the context of sojourns across short or intermediate time intervals (Greischel et al., 2016; Niehoff et al., 2017; Zimmermann & Neyer, 2013). Thus, the question arises as to how sojourners develop in the long run compared to their fellow students. Put differently, we ask whether and to what extent we can observe persistent differences between the trait change trajectories of sojourners and their peers who never went abroad (i.e., *stayers*). Using a longitudinal (nonrandom) control group design, we investigated three possible patterns, i.e., reversed, accentuated or sustained sojourn effects.

According to a dynamic equilibrium (or classic set-point) model, changes in individual characteristics are time-limited, and reflect fluctuating experiences (Headey & Wearing, 1989; Lykken & Tellegen, 1996). From this perspective, personality trait change differences between sojourners and stayers are expected to reverse after return from the stay abroad in the long run (*reversed effect*). Let us assume that Tina is a German student that had gone abroad for one semester. During her sojourn, she made multiple new experiences, for example, regarding food choices, habits, and ways of life. Moreover, she engaged with people from several countries with different backgrounds than her own. These daily experiences resulted in changes in Tina's daily behavior. For example, Tina found herself with a lot of new behaviors and ideas due to activities with friends from other cultural backgrounds. At the same time, her friends from home who did not study abroad did not make such new experiences. Over time, these environmental state differences between Tina and her friends solidified and thus promoted trait change differences in openness between them (see Roberts, 2018; Roberts & Jackson, 2008). After Tina has left her host country and returned to her home context, the experiential differences between Tina and her friends at home diminish. In case that openness normatively decreases (or increases), Tina's openness decreases more steeply (or increases less steeply) than her friends', leading to decreased mean-level differences between them after her return as compared to the previous measurement occasion(s) (see Figure 1). One reason for a reversed effect might be that Tina picks up old

habits and behaviors, and adapts again to her home's norms (Denissen, van Aken, Penke, & Wood, 2013).[5]

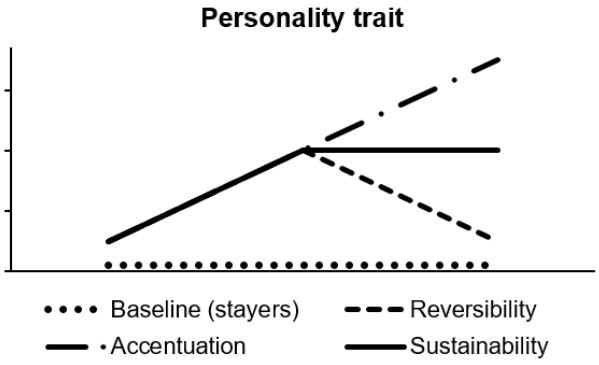

Figure 1. Examples of long-term trait change patterns due to sojourn experiences in comparison to stayers' baseline changes.

In contrast to reversing differences, we could also observe accentuating differences between sojourners and stayers in the long run (see also Jackson, Thoemmes, Jonkmann, Lüdtke, & Trautwein, 2012), possibly due to mutual reinforcement between corresponsive trait levels and life experiences (Jeronimus, Riese, Sanderman, & Ormel, 2014; Le, Donnellan, & Conger, 2014; Roberts, Caspi, & Moffitt, 2003). For example, it has recently been shown that previous sojourns can predict future sojourns (Netz & Jaksztat, 2014; Niehoff et al., 2017). This is an example of cumulative differences in terms of recurring life events that may eventually act to accentuate differences between sojourners and stayers (*accentuated effect*). Let us imagine that Tina comes back home. Although she had looked forward to meeting all her friends, she also misses the experiences she made in a different country, and decides to go abroad again. Meeting new people and making new experiences again reinforce her new habits and behaviors. She feels good this way, and decides to go abroad for longer time periods on a regular basis now. While her friends at home decrease in openness slowly but incrementally (see, for example, McCrae et al., 1999) Tina's openness may not decrease or even increases, leading to accentuated differences between her and her friends' openness over time (see Figure 1).

From a third perspective, meaningful experiences due to sojourns could result in stable trait differences between sojourners and stayers, being evident across years (*sustained effect*). One explanation for such a pattern might be that sojourners adopt new reference values during

[5] Figure 1 suggests that stayers do not show any trait changes. In fact, this is highly unlikely (Graham et al., 2020). In addition, we want to stress that a reversed effect does not necessarily imply the same mean levels in sojourners at the start and at the end of the study. Rather, it means that developmental trends turn back, leading to decreasing differences between sojourners and stayers after return.

their stay abroad, which they still hold some years later (Denissen et al., 2013). Tina would, for example, return home a bit more open to new experiences compared to before with an enhanced difference to her friends at home. After her return, she might show the same developmental trends as her friends (e.g., the magnitude of their openness decreases does not differ), but her openness persists at a comparably higher trait level (compare Figure 1).

The Present Study

With the PEDES study design, we compared sojourners' trait changes across a period of five years to fellow students of the same cohort who had not lived abroad. We examined three time points around a sojourn: T1 (directly before the start of the sojourn), T2 (approximately six months after sojourn start), and T3 (approximately nine months after sojourn start). Furthermore, unlike other studies, we conducted a follow-up measurement occasion about five years after sojourn start (T4). This fourth measurement occasion extended previous studies of short-term sojourn effects on personality trait changes, and enabled us to examine the long-term effects of sojourn experiences.

Change in an individual's social and cultural environment for several months – as during a sojourn – could come along with any of the three plausible scenarios discussed in the preceding section. We are not aware of any study that has examined long-term trait changes of sojourners. Thus, we had no a-priori hypotheses. The study was not preregistered. By analyzing the five-year latent trait changes of students with and without sojourn experiences, we explored which of the long-term sojourn effects (reversed, accentuated, or sustained) were and were not supported by the data. Sojourners and stayers of the same sample had already been found to differ in their short-term personality trait changes (Zimmermann & Neyer, 2013). To control for these effects, we included the previous measurement intervals (T1–T3) in our analysis models.

Method

Participants and Procedure

Data come from the PEDES longitudinal study (for further details regarding recruitment methods and selection criteria, see Zimmermann and Neyer, 2013). Data to this study can be drawn from the open science framework (https://osf.io/pmy57/).[6] The current

[6] We intend to provide open access to all data of the PEDES project once all planned project publications have been completed.

investigation went beyond the previous study by adding a follow-up measurement occasion (T4) about 5 years after the first measurement took place. All measurements were carried out using online questionnaires. During the first study period, a sample of N = 1,134 German university students were repeatedly tracked over the course of the academic year 2009/10 (T1, T2, T3). Data were collected two weeks before (T1), six months after (T2; time range: 4–8 months), and nine months after (T3; time range: 6–13 months) sojourners' individual dates of departure. Control students who did not live abroad during that time completed the same questionnaires at comparable time points during the academic year 2009/10.

To compare the long-term personality trait changes of control students (here, stayers) and sojourners, participants were again surveyed about five years after T1. The T4 measurements were taken between December 2014 and March 2015. The primary interest of the present study was to compare participants' trait changes during the follow-up period of about 4 years (T3–T4). To capture sojourn effects on personality trait changes in the long run under the control of earlier effects, we included initial trait levels at T1 and participants' trait changes across T1, T2, and T3 in the present analyses.

From the initial sample (N = 1,134), we excluded n = 39 participants who had reported their main place of residence at T4 outside of Germany. From the resulting total sample of n = 1,095 students, n = 498 were *sojourners* who went abroad shortly after the start of the study at T1. The vast majority of the sojourners moved to a European country (84.9%, n = 423), 6.0% (n = 30) to Asia, 5.4% (n = 27) to the United States or to Canada, 0.6% (n = 3) to South America, and 0.2% (n = 1) to Australia. Moreover, 1.4% (n = 7) participants went to Turkey, and 0.8% (n = 4) to Russia. Three participants did not provide an answer to this question. We defined sojourners that had spent 20–32 weeks during the academic year 2009/10 in their host countries as short-term sojourners (n = 218), and those who spent more than 32 weeks as long-term sojourners (n = 280). Stayers were students that did not live abroad before or during the academic year 2009/10 (n = 597). Mean age at survey start was 22.59 years (SD = 2.60, age range: 18–41 years), and 78.0% of the participants were females (n = 854).

Drop-out. At T3, n = 1,059 (96.7%) of our sample agreed to be contacted again in the future for a possible follow-up. At T4, however, only n = 441 participants (40.3% of the original sample) provided information on their Big Five traits. As we had not communicated concrete plans for a follow-up study, we assume that most participants saw their contribution as completed after T3. Moreover, during the four-year time interval between T3 and T4, we

did not implement any measures of panel maintenance. Against this background, and compared to other panel studies, a retention rate of 40.3% can be seen as acceptable. To test whether drop-out patterns were completely at random across the five-year time span, especially between T3 and T4, and for each variable investigated, we computed Missing-Completely-at-Random (MCAR) test statistics. Little's MCAR test indicated no association between sojourn status, age, sex, or trait level at any time point between T1 and T3, and later drop-out, $\chi^2(16) = 9.219$, $p = .904$. Of all participants at T4, $n = 372$ had graduated in the meantime. $N = 33$ participants indicated that they were studying in an undergraduate degree program, and $n = 67$ participants in a postgraduate course. $N = 235$ participants indicated that they were employed in a regular job (no apprenticeship). Other possible options were, for example, self-employment, teacher training, parental leave, etc.[7]

Measures

Personality trait measures. We applied the German version of the Big Five Inventory (BFI; Lang, Lüdtke, & Asendorpf, 2001) at all four measurement occasions. Participants answered the 42 items on a five-point Likert scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). Internal consistency coefficients of the personality trait measures can be found in Table 1. Moreover, Table 1 contains zero-order correlations of all Big Five personality traits, including rank-order stabilities for the total sample and subsamples. Rank-order stabilities across measurement occasions were high for all traits. Table 2 shows means and standard deviations for each measurement occasion, and effect sizes of manifest mean-level differences in repeated measures for both the full sample and subgroups, standardized on the standard deviation of the respective first measurement time point.

Sojourn status. For the analyses, two different kinds of status definitions were used. First, in the main (two-group) analyses, we differentiated between participants without any sojourn experiences before or during the academic year 2009/10 (*stayers*) and participants

[7] Please note that multiple references were possible. Here, we do not provide information on the full range of answering options. Please refer to the codebook for further information (https://osf.io/pmy57/).

Table 1

Zero-Order Correlations and Internal Consistency Coefficients of all Big Five Trait Measures in the Total Sample and in the Respective Groups

	T1					T2				
	O	C	E	A	N	O	C	E	A	N
T1										
O	*.83*	**.10**/ .08	**.28**/ **.24**	.06/ .01	**–.10**/–.03	**.86**/ **.85**	**.10**/ .03	**.26**/ **.23**	**.07**/ .03	**–.08**/–.01
C	**.11**/ .03/ **.19**	*.82*	**.20**/ **.17**	**.32**/ **.31**	**–.16**/**–.14**	**.08**/ .05	**.85**/ **.87**	**.21**/ **.19**	**.27**/ **.26**	**–.14**/**–.11**
E	**.29**/ **.27**/ **.31**	**.19**/ .11/ **.26**	*.88*	**.17**/ **.17**	**–.32**/**–.30**	**.25**/ **.20**	**.21**/ **.17**	**.87**/ **.88**	**.18**/ **.16**	**–.30**/**–.26**
A	**.09**/ .02/ **.16**	**.32**/ **.25**/ **.36**	**.16**/ .02/ **.26**	*.72*	**–.28**/**–.28**	**.09**/ .04	**.30**/ **.30**	**.20**/ **.18**	**.78**/ **.80**	**–.26**/**–.23**
N	**–.13**/**–.18**/–.10	**–.16**/**–.15**/**–.16**	**–.30**/**–.34**/**–.27**	**–.26**/**–.24**/**–.28**	*.82*	**–.09**/–.02	**–.18**/**–.15**	**–.31**/**–.29**	**–.30**/**–.29**	**.80**/ **.80**
T2										
O	**.87**/ **.85**/ **.89**	**.10**/–.01/ **.17**	**.27**/ **.24**/ **.29**	**.12**/ .04/ **.18**	**–.11**/–.10/**–.13**	*.84*	**.12**/ .06	**.29**/ **.24**	**.12**/ .07	**–.10**/–.02
C	**.15**/ .06/ **.23**	**.84**/ **.82**/ **.84**	**.22**/ **.20**/ **.24**	**.29**/ **.23**/ **.33**	**–.18**/**–.16**/**–.19**	**.17**/ .04/ **.26**	*.83*	**.25**/ **.21**	**.33**/ **.31**	**–.20**/**–.17**
E	**.27**/ **.26**/ **.29**	**.21**/ .12/ **.28**	**.85**/ **.86**/ **.84**	**.21**/ .06/ **.31**	**–.29**/**–.34**/**–.25**	**.30**/ **.31**/ **.31**	**.27**/ **.23**/ **.29**	*.89*	**.26**/ **.21**	**–.38**/**–.33**
A	.08/–.04/ **.17**	**.26**/ **.16**/ **.33**	**.16**/ .06/ **.23**	**.75**/ **.71**/ **.78**	**–.26**/**–.20**/**–.31**	**.15**/ .07/ **.21**	**.32**/ **.27**/ **.36**	**.28**/ **.23**/ **.33**	*.72*	**–.37**/**–.33**
N	**–.13**/–.11/**–.15**	**–.14**/**–.18**/–.11	**–.30**/**–.29**/**–.30**	**–.29**/**–.20**/**–.34**	**.79**/ **.79**/ **.78**	**–.13**/–.06/**–.17**	**–.21**/**–.22**/**–.19**	**–.40**/**–.41**/**–.40**	**–.39**/**–.32**/**–.44**	*.84*
T3										
O	**.86**/ **.86**/ **.85**	.08/–.03/ **.17**	**.23**/ **.19**/ **.26**	**.13**/ .11/ **.15**	**–.12**/**–.15**/–.11	**.89**/ **.89**/ **.90**	**.14**/ .03/ **.23**	**.27**/ **.26**/ **.29**	**.15**/ .10/ **.19**	**–.14**/–.12/**–.15**
C	**.14**/ .07/ **.22**	**.84**/ **.84**/ **.83**	**.22**/ **.15**/ **.27**	**.33**/ **.28**/ **.36**	**–.18**/**–.16**/**–.19**	**.17**/ .06/ **.25**	**.86**/ **.86**/ **.86**	**.27**/ **.20**/ **.32**	**.32**/ **.26**/ **.36**	**–.20**/**–.22**/**–.17**
E	**.26**/ **.24**/ **.28**	**.20**/ .11/ **.26**	**.85**/ **.85**/ **.85**	**.21**/ .09/ **.29**	**–.30**/**–.33**/**–.27**	**.28**/ **.25**/ **.31**	**.28**/ **.22**/ **.30**	**.90**/ **.88**/ **.91**	**.27**/ **.21**/ **.30**	**–.36**/**–.36**/**–.36**
A	.06/–.06/ **.14**	**.21**/ **.15**/ **.25**	**.13**/ .02/ **.21**	**.77**/ **.80**/ **.76**	**–.23**/**–.14**/**–.30**	**.11**/–.01/ **.20**	**.26**/ **.19**/ **.30**	**.25**/ **.15**/ **.33**	**.82**/ **.80**/ **.83**	**–.32**/**–.22**/**–.39**
N	**–.13**/–.11/**–.15**	**–.10**/–.12/–.08	**–.27**/**–.23**/**–.29**	**–.27**/**–.26**/**–.29**	**.78**/ **.77**/ **.79**	**–.14**/–.05/**–.20**	**–.15**/**–.16**/**–.16**	**–.35**/**–.33**/**–.36**	**–.35**/**–.31**/**–.38**	**.84**/ **.83**/ **.86**
T4										
O	**.70**/ **.70**/ **.70**	–.07/–.16/ .02	.11/ .17/ .05	–.04/–.08/ .01	.03/–.03/ .06	**.75**/ **.81**/ **.71**	–.06/–.10/–.03	.10/ .17/ .03	.02/ .04/ .01	.07/ .09/ .06
C	–.04/–.02/–.04	**.70**/ **.74**/ **.66**	**.20**/ .14/ **.26**	**.18**/ **.23**/ .13	–.05/–.04/–.01	–.05/–.04/–.04	**.70**/ **.74**/ **.65**	**.25**/ **.32**/ **.19**	**.16**/ .21/ .13	–.09/**–.22**/ .04
E	**.16**/ .20/ .13	.14/ .14/ .14	**.81**/ **.81**/ **.81**	.10/ .15/ .06	**–.20**/–.22/–.18	**.16**/ **.21**/ .13	**.18**/ .22/ .15	**.76**/ **.79**/ **.74**	.11/ .14/ .08	**–.22**/–.10/**–.29**
A	–.00/–.11/ .10	**.26**/ .21/ **.30**	.06/–.08/ **.18**	**.63**/ **.66**/ **.60**	**–.15**/–.06/**–.20**	.07/ .00/ .13	**.27**/ .20/ **.33**	.13/ .04/ **.22**	**.65**/ **.66**/ **.64**	**–.25**/–.09/**–.36**
N	–.03/–.05/–.05	–.02/ .03/–.02	–.10/ .03/**–.21**	**–.29**/**–.28**/**–.29**	**.69**/ **.61**/ **.73**	–.02/ .05/–.09	–.03/ .05/–.04	–.13/–.04/**–.20**	**–.32**/**–.25**/**–.37**	**.70**/ **.63**/ **.73**

Note. Cases with missing values on one or more items were excluded. O = openness, C = conscientiousness, E = extraversion, A = agreeableness, N = neuroticism. Internal consistency coefficients based on the total sample are printed in italic along the diagonal.

Above the diagonal: Total sample [a]/stayers [b]. *Below the diagonal:* All sojourners [c]/short-term sojourners [d]/long-term sojourners [e].

[a] $N_{T1–T3}$ = 1,095, n_{T4} = 441. [b] $N_{T1–T3}$ = 597, n_{T4} = 239. [c] N_{T1-T3} = 498, N_{T4} = 202. [d] $N_{T1–T3}$ = 218, n_{T4} = 88. [e] $N_{T1–T3}$ = 280, n_{T4} = 114.

Significant correlations ($p < .05$) are printed in boldface.

Table 1 *continued*

T3					T4				
O	C	E	A	N	O	C	E	A	N
.84/ .82	**.10**/ .04	**.26/ .24**	.06/ .03	**–.09**/–.01	**.71/ .72**	–.02/–.02	**.18/ .18**	.04/ .06	–.01/ .03
.07/ .04	**.85/ .86**	**.21/ .20**	**.26/ .28**	**–.11/–.10**	.02/ .05	**.75/ .77**	**.16/ .14**	**.26/ .23**	–.06/–.07
.23/ .20	**.22/ .19**	**.86/ .87**	**.17/ .16**	**–.28/–.24**	**.13**/ .11	**.21/ .18**	**.78/ .74**	**.13/ .15**	**–.16/–.18**
.07/ .01	**.31/ .29**	**.20/ .18**	**.78/ .78**	**–.25/–.23**	.02/ .03	**.24/ .26**	**.14/ .13**	**.66/ .67**	**–.29/–.28**
–.09/–.02	**–.19/–.18**	**–.32/–.31**	**–.28/–.30**	**.80/ .81**	.00/ .02	**–.10**/–.10	**–.22/–.19**	**–.21/–.23**	**.69/ .68**
.88/ .87	**.12**/ .05	**.27/ .22**	**.10**/ .06	**–.10**/–.02	**.74/ .73**	–.00/ .00	**.16/ .14**	.07/ .05	.00/ .05
.09/ .03	**.87/ .88**	**.25/ .21**	**.29/ .30**	**–.16/–.14**	.01/ .04	**.75/ .78**	**.18/ .14**	**.28/ .27**	–.09/–.11
.27/ .23	**.26/ .23**	**.90/ .90**	**.23/ .19**	**–.33/–.28**	**.14/ .14**	**.23/ .19**	**.77/ .76**	**.17/ .17**	**–.18/–.18**
.10/ .03	**.32/ .30**	**.24/ .20**	**.83/ .83**	**–.34/–.31**	.03/ .01	**.21/ .23**	**.12**/ .09	**.68/ .69**	**–.31/–.28**
–.09/–.02	**–.21/–.19**	**–.36/–.32**	**–.32/–.29**	**.85/ .85**	.01/ .00	**–.12**/–.10	**–.25/–.22**	**–.28/–.27**	**.73/ .75**
.84	**.12**/ .05	**.29/ .26**	**.10**/ .04	**–.12**/–.04	**.73/ .70**	.01/–.01	**.19/ .16**	.07/ .02	.02/ .09
.18/ .06/ **.27**	*.83*	**.28/ .25**	**.34/ .33**	**–.21/–.20**	.02/ .05	**.75/ .79**	**.19/ .17**	**.30/ .29**	**–.12/–.15**
.29/ .26/ .33	**.29/ .20/ .34**	*.89*	**.26/ .23**	**–.36/–.33**	**.15/ .17**	**.24/ .23**	**.79/ .76**	**.16/ .17**	**–.20/–.20**
.15/ .08/ **.20**	**.33/ .28/ .37**	**.27/ .18/ .33**	*.74*	**–.34/–.31**	.06/ .06	**.25/ .29**	**.17/ .14**	**.71/ .72**	**–.33/–.30**
–.17/–.14/–.19	**–.19/–.20/–.19**	**–.37/–.36/–.38**	**–.35/–.28/–.40**	*.84*	.01/–.00	**–.11**/–.11	**–.23/–.19**	**–.27/–.28**	**.71/ .71**
.75/ .79/ .73	–.05/–.11/ .01	.09/ .18/ .03	.02/–.01/ .04	.07/ .04/ .09	*.83*	.07/ .07	**.23/ .21**	**.14/ .15**	–.02/ .02
–.01/–.02/ .02	**.68/ .77/ .59**	**.20/ .23**/ .17	**.16**/ .19/ .14	–.06/–.20/ .06	.04/–.07/ .12	*.80*	**.27/ .24**	**.30/ .30**	**–.13/–.17**
.19/ .22/ .16	**.18**/ .17/ **.19**	**.81/ .87/ .77**	**.14**/ .19/ .11	**–.22/–.22/–.22**	**.22/ .29**/ .18	**.26/ .24/ .27**	*.88*	**.20/ .19**	**–.26/–.26**
.09/ .08/ .12	**.29/ .24/ .34**	.11/ .01/ .18	**.67/ .70/ .66**	**–.22**/–.14/**–.28**	.10/–.01/ **.20**	**.27**/ .18/ **.33**	**.17**/ .13/ **.20**	*.72*	**–.35/–.34**
–.03/–.06/–.03	–.04/ .04/–.07	**–.16**/–.07/**–.20**	**–.33/–.28/–.38**	**.69/ .63/ .73**	–.02/ .02/–.05	–.04/–.01/–.02	**–.22**/–.08/**–.31**	**–.35/–.22/–.45**	*.85*

Table 2
Group-Based Descriptive Statistics

	M (*SD*)				Effect size for repeated measures [a]			
Group	T1	T2	T3	T4	d_{12}	d_{23}	d_{34}	d_{14}
Full sample [b]								
O	3.67 (0.64)	3.65 (0.62)	3.65 (0.62)	3.59 (0.64)	–0.03	0.00	–0.10	–0.13
C	3.58 (0.63)	3.55 (0.62)	3.53 (0.61)	3.64 (0.59)	–0.05	–0.03	0.18	0.10
E	3.44 (0.74)	3.40 (0.74)	3.44 (0.73)	3.35 (0.73)	–0.05	0.05	–0.12	–0.12
A	3.63 (0.55)	3.62 (0.54)	3.62 (0.54)	3.66 (0.54)	–0.02	0.00	0.07	0.05
N	2.97 (0.71)	2.95 (0.71)	2.89 (0.71)	2.91 (0.74)	–0.03	–0.08	0.03	–0.08
Stayers [c]								
O	3.59 (0.63)	3.56 (0.61)	3.57 (0.61)	3.52 (0.61)	–0.05	0.02	–0.08	–0.11
C	3.52 (0.65)	3.48 (0.61)	3.47 (0.62)	3.56 (0.60)	–0.06	–0.02	0.15	0.06
E	3.32 (0.73)	3.29 (0.73)	3.34 (0.73)	3.24 (0.70)	–0.04	0.07	–0.14	–0.11
A	3.60 (0.57)	3.56 (0.54)	3.56 (0.54)	3.61 (0.55)	–0.07	0.00	0.09	0.02
N	3.07 (0.69)	3.07 (0.71)	3.01 (0.71)	3.00 (0.75)	0.00	–0.08	–0.01	–0.10
All sojourners [d]								
O	3.76 (0.64)	3.75 (0.62)	3.75 (0.62)	3.67 (0.65)	–0.02	0.00	–0.13	–0.14
C	3.65 (0.60)	3.62 (0.62)	3.61 (0.59)	3.73 (0.55)	–0.05	–0.02	0.20	0.13
E	3.59 (0.72)	3.54 (0.73)	3.57 (0.71)	3.49 (0.73)	–0.07	0.04	–0.11	–0.14
A	3.67 (0.53)	3.68 (0.53)	3.69 (0.53)	3.72 (0.52)	0.02	0.02	0.06	0.09
N	2.86 (0.72)	2.81 (0.68)	2.76 (0.68)	2.80 (0.73)	–0.07	–0.07	0.06	–0.08
Short-term sojourners [e]								
O	3.71 (0.62)	3.74 (0.60)	3.71 (0.60)	3.66 (0.66)	0.05	–0.05	–0.08	–0.08
C	3.73 (0.58)	3.70 (0.60)	3.70 (0.57)	3.84 (0.53)	–0.05	0.00	0.25	0.19
E	3.60 (0.73)	3.60 (0.76)	3.64 (0.70)	3.51 (0.72)	0.00	0.05	–0.19	–0.12
A	3.70 (0.51)	3.71 (0.52)	3.69 (0.50)	3.76 (0.55)	0.02	–0.04	0.14	0.12
N	2.81 (0.73)	2.77 (0.67)	2.78 (0.70)	2.65 (0.67)	–0.05	0.01	–0.19	–0.22
Long-term sojourners [f]								
O	3.79 (0.64)	3.77 (0.64)	3.77 (0.63)	3.68 (0.65)	–0.03	0.00	–0.14	–0.17
C	3.59 (0.60)	3.55 (0.63)	3.54 (0.59)	3.64 (0.55)	–0.07	–0.02	0.17	0.08
E	3.58 (0.72)	3.49 (0.71)	3.51 (0.73)	3.47 (0.74)	–0.13	0.03	–0.05	–0.15
A	3.65 (0.55)	3.67 (0.54)	3.68 (0.56)	3.69 (0.49)	0.04	0.02	0.02	0.07
N	2.90 (0.71)	2.84 (0.69)	2.75 (0.66)	2.91 (0.75)	–0.08	–0.13	0.24	0.01

Note. O = openness, C = conscientiousness, E = extraversion, A = agreeableness, N = neuroticism.
[a] Within-group repeated-measures raw-score metric (compare Morris & DeShon, 2002). More specifically, we calculated *d* as the difference between the mean levels of two time points divided by the standard deviation of the respective first time point.
[b] $N_{T1–T3}$ = 1,095; n_{T4} = 441.
[c] $N_{T1–T3}$ = 597, n_{T4} = 239.
[d] $N_{T1–T3}$ = 498, n_{T4} = 202.
[e] $N_{T1–T3}$ = 218, n_{T4} = 88.
[f] $N_{T1–T3}$ = 280, n_{T4} = 114.

with sojourn experiences during the academic year 2009/10 (*sojourners*). Based on this information, we constructed the dummy variable sojourn T1, which was coded 0 (*no sojourn*) and 1 (*sojourn*). At T2 and T3, sojourners were asked about their current country of residence (in Germany or abroad). This allowed us to additionally explore differences in sojourn effects depending upon sojourn duration. *Short-term sojourners* (one semester abroad) had already

returned to Germany by T3, while *long-term sojourners* (at least two semesters abroad) still lived in their host country at T3. In the more extensive (three-group) analyses, we thus used two dummy variables to differentiate between these sojourn groups. The first variable, called short-term sojourn, was coded 0 (*no short-term sojourn*) and 1 (*short-term sojourn*). It is important to note that coding 0 included all participants that had not done a short-term sojourn, i.e., stayers and long-term sojourners. The second variable was called long-term sojourn, and was coded 0 (*no long-term sojourn*) and 1 (*long-term sojourn*). All participants who had not done a long-term sojourn, i.e., stayers and short-term sojourners, were coded 0 in this variable.

Covariates. To control for possible confounding effects, we included age, sex, and the number of sojourns between T3 and T4 as covariates in the analyses. Sex was coded 0 (*male*) and 1 (*female*). To yield information about additional sojourn experiences after the end of the first study period in 2009/10, we asked participants at T4 how often they had lived abroad for more than two months since T3. Of n = 443 participants who answered this question, n = 101 participants indicated that they had lived abroad at least once, while n = 342 indicated that they have never lived abroad since T3 (M = 0.32, SD = .71, range: 0–6). Only n = 29 of the former stayers went abroad between T3–T4 (*novice sojourners*).

Analytical Strategy

We examined our research questions using multiple approaches. We will shortly summarize them here, before explaining the methodological features of our main analyses. Preliminary latent trait change analyses only included time points T1 and T4. We then investigated sojourners' and stayers' latent trait changes more deeply by also including T2 and T3 in latent neighbor-change models. For both approaches, were carried out univariate as well as multivariate analyses. Although we examined results based on listwise deletion and FIML to handle missing data, we will mostly refer to the analyses based on FIML. To compare short-term and long-term sojourns regarding their effects, we additionally run analyses separating sojourners according to the duration of their stay.

In the following sections, we describe the latent neighbor-change modeling approach in more detail. We first reflect on the latent variable and latent change modeling, followed by the characteristics of the multivariate approach. One part of this section is dedicated to describing the estimation of sojourn effects. All input and output files are available via the

open science framework (https://osf.io/pmy57/). Both the main and additional analyses were carried out using Mplus version 8 (Muthén & Muthén, 1998–2017).

Latent variable modeling. We modeled individual personality trait levels as latent variables to control for measurement error (Steyer, Mayer, Geiser, & Cole, 2015). To that end, we adopted the same method as Zimmermann and Neyer (2013): Items were assigned to respectively two parcels per personality trait on the basis of the item-to-construct method (Little, Cunningham, Shahar, & Widaman, 2002). The goal of this method is to receive balanced parcels in terms of factor loadings. Internal consistency coefficients of the parcels can be found in the supplementary Table S1. For the latent analyses, all latent factor indicators (i.e., parcels) were standardized.

Latent change modeling. We captured personality trait changes within a five-year time span using latent change models (McArdle & Hamagami, 2002; Steyer, Eid, & Schwenkmezger, 1997). For a preliminary check of the sustainability of sojourn effects, we analyzed personality trait changes across T1–T4 only. Thereafter, we analyzed all four time points using latent neighbor-change models to capture (discontinuous) change during each time interval (i.e., from T1–T2, from T2–T3, and from T3–T4; Steyer, Partchev, & Shanahan, 2000). In latent neighbor-change models, latent trait change reflects the difference between latent trait levels at two neighboring measurement occasions controlling for measurement error (Steyer et al., 1997; Steyer et al., 2000). To interpret latent variable change as trait change, the latent variables must have the same meaning across time points (Muthén & Muthén, 1998–2017). Therefore, we implemented strict measurement invariance across measurement occasions by equating the latent variable structure, parcel loadings, intercepts, and residual variances of the same parcels per trait across time (Marcusson-Clavertz & Kjell, 2018).[8]

We established an indicator-specific factor for the second parcel in all models to account for common variance in the manifest variables across time that was not captured by the latent variables (see Geiser & Lockhart, 2012, for a comparison of this approach to other

[8] Residual variances do not need to be equal across time points to interpret the meaning of their underlying latent variables (van de Schoot, Lugtig, & Hox, 2012). Formal model comparisons yielded significantly worse model fits of the more restrictive models with equated residuals in comparison to the less restrictive models in some cases. Still, we decided to equate residual variances across time for all models. We did this as the more restrictive models with strict measurement invariance also revealed good overall model fits that were comparable to those of the less restrictive models with scalar measurement invariance (see supplementary Table S2). Results were robust across both variants of measurement invariance. Input and output files of all analyses can be drawn from https://osf.io/pmy57/.

procedures). Loadings of the respective parcels on the indicator-specific factor were modeled to be time-invariant. This factor was not allowed to correlate with the latent trait factors of the same construct (Reuter et al., 2010). Latent change variables of the same trait across different time intervals were allowed to correlate with each other as change within one trait across a specific time interval was much likely not independent from changes within the same trait across other time intervals. That is, trait change in openness between T1 and T2 was allowed to correlate with T2–T3 and T3–T4 trait changes in openness, and trait change in openness between T2 and T3 was allowed to correlate with T3–T4 trait change in openness. Figure 2 provides a conceptual overview of how we assessed the effects of a sojourn during the academic year 2009/10 on trait changes between T1 and T2, T2 and T3, and between T3 and T4.

Multivariate trait change modeling. Although Figure 2 provides an exemplary model for only one trait, we also included all Big Five traits in multivariate models to compare univariate and multivariate findings. In line with the analytical strategy used by Zimmermann and Neyer (2013), we allowed correlations between trait levels at T1 across all traits to account for the fact that operationalizations of personality traits are empirically not perfectly independent of each other (e.g., Lang et al., 2001). In addition, all trait levels at T1 were allowed to correlate with T1–T2, T2–T3, and T3–T4 trait change variables. While latent change variables across different traits were also allowed to correlate with each other within the same time interval, we did not allow cross-trait cross-interval correlations for the sake of parsimony. Finally, the residual factors were not allowed to correlate with each other, but we did not constrain their correlations with trait factors of other constructs (Reuter et al., 2010).[9]

Sojourn effects modeling. By including sojourn effects on earlier trait changes (within T1–T2 and T2–T3 intervals), the model tested the effects of a sojourn during the academic year 2009/10 on trait changes between T3 and T4 after controlling for prior effects

[9] Allowing the method factors to correlate with each other did not alter the pattern of results.

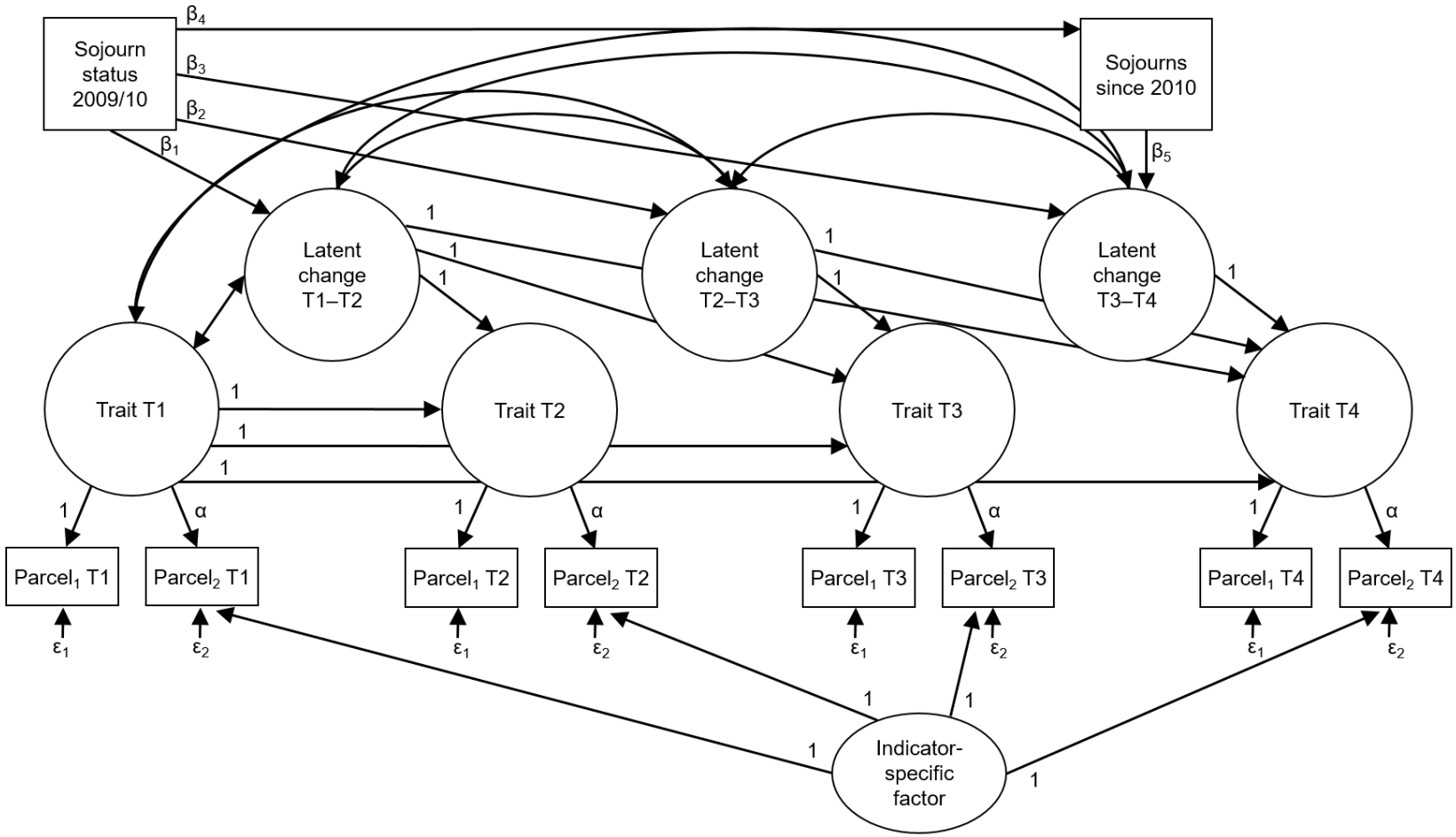

Figure 2. Conceptual model of sojourn effects on trait changes across T1–T2, T2–T3, and T3–T4. For reasons of parsimony, we only printed an exemplary univariate latent neighbor-change model, although we also imputed multivariate latent neighbor-change models. We controlled for age, sex, and sojourns since T3 as possible confounds. A time-invariant indicator-specific method factor was constructed per trait to account for trait-independent variance that was related to the parcels. Residuals were equated across time.

on T1–T2 and T2–T3 changes (see Figure 2). With this approach, it was also possible to explore the timing of sojourn effects. Specifically, we extended the previous analyses reported by Zimmermann and Neyer (2013) by investigating whether early change ($\Delta_{T1–T2} > \Delta_{T2–T3}$), continuous change ($\Delta_{T1–T2} = \Delta_{T2–T3}$), or late change ($\Delta_{T1–T2} < \Delta_{T2–T3}$) occurred.

Differentiating between sojourn groups. Besides testing the effect of a sojourn in itself in the first set of analyses (two-group analyses), we also compared the effects of short-term and long-term sojourns regarding the magnitude of change and the timing of their effects in additional three-group analyses. For example, it has been shown that adaptation processes during a sojourn are dependent on the time spent abroad (see Ward, Okura, Kennedy, & Kojima, 1998). Moreover, the transition back home might itself have an effect on trait changes (compare Christofi & Thompson, 2007), which can also depend on sojourn duration (Tamura & Furnham, 1993). At T3, short-term sojourners had already returned to their home country, whereas long-term sojourners had not. This resulted in a different meaning of the sojourn status variables on T2–T3 trait changes for short-term and long-term sojourners: the transition back home and prolonged sojourn, respectively. To consider these group differences, we regressed the latent trait change variables on two dummy-coded sojourn variables: short-term sojourn and long-term sojourn. All input and output files of these three-group (stayers, short term sojourners, and long-term sojourners) analyses are available via the open science framework (https://osf.io/pmy57/).

Covariates. As covariates, we included sex, age, and the number of sojourns since T3. We regressed all trait-change variables on age and sex. T3–T4 trait change was also regressed on the number of sojourns between T3 and T4. This analysis tests the effect of the 2009/10 sojourn on later trait change between T3 and T4 while controlling for the effect of having lived abroad since T3. In addition, we regressed the number of sojourns since T3 on sojourn status at T1 to account for potential associations between previous sojourns during the academic year 2009/10 on further sojourn experiences during the T3–T4 interval.

Analytical criteria. We evaluated the absolute goodness of model fit using the RMSEA (root mean square error of approximation), CFI (comparative fit index), and SRMR (standardized root mean square residual). RMSEA portrays the average proportion of misfit per degree of freedom, CFI is an indicator for the relative non-centrality between a hypothesized model and the null model of the modified independency, and SRMR refers to the squared root of the average squared residuals. While for CFI, a higher value indicates a

better fit, the contrary is true for RMSEA and SRMR. Hu and Bentler (1999) have described criteria for a good model fit with RMSEA ≤ .06, CFI ≥ .95, and SRMR ≤ .08.

Parameter estimations were based on a Maximum Likelihood procedure with robust estimation of standard errors. The latent change models were fitted using the Full Information Maximum Likelihood (FIML) method, which is a model-based approach to handle missing data (Schafer & Graham, 2002). In comparison to other missing data handling procedures, it allows for more precise parameter estimation and retains statistical power because no observations are deleted (Enders, 2011). To scrutinize the robustness of our findings from the latent trait-change models, we also analyzed descriptive mean-level trends and carried out model comparisons based on Bayesian information criteria.

Results

Manifest Mean-Level Trends

Figure 3 illustrates descriptive mean-level trends and standard errors per measurement occasion based on the raw means of personality trait variables for stayers and sojourners (i.e., the unweighted sample of short-term and long-term sojourners; see also Table 2). The figure indicated initial selection effects for all traits. Compared to stayers, sojourners tended to show higher levels in openness, conscientiousness, extraversion, agreeableness, and emotional stability (vs. neuroticism). Moreover, these differences tended to be constant across all four measurement occasions. See also supplementary Figure S1, which contains raw means and standard errors of stayers, short-term sojourners, and long-term sojourners per measurement occasion separately.

Figures 3 and S1 also show that if there were significant sojourn effects with increased differences between sojourners and stayers during a sojourn (T1–T2 for short-term sojourners and T1–T3 for long-term sojourners), these effects must have been very small and mostly diminished until T4 as indicated by rather parallel linear trends. However, manifest personality trait score trajectories were confounded with measurement error and potentially biased due to unconsidered inter-trait correlations and covariates (age, sex, and repeated sojourns). This could lead to reduced accuracy of measurements and obscure or bias sojourn effects. Therefore, we ran univariate and multivariate latent neighbor-change models controlling for potential biases. Please refer to supplementary Figure S2 for time-specific boxplots of the three groups based on listwise deletion.

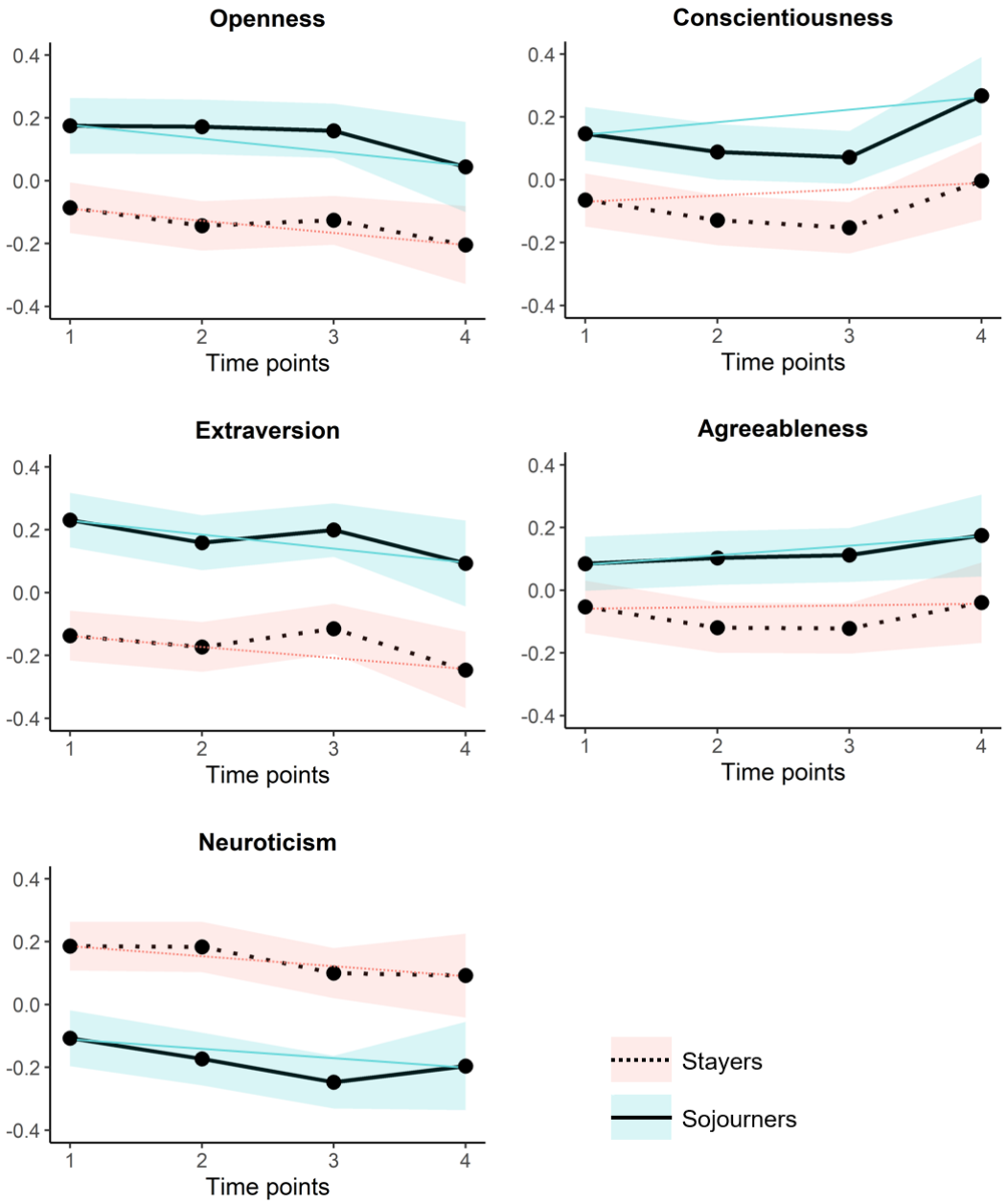

Figure 3. Descriptive mean-level differences of stayers and sojourners based on raw means of manifest personality trait scores. We plotted the raw means standardized around the grand mean of each trait to make the figures more comparable. Estimates include group-based 95% confidence intervals to illustrate measurement inaccuracy. We included a linear trend line between T1 and T4 in each group's respective color theme to make general linear change trends for the whole study period more easily comparable across groups. Please note that trends appear much steeper for the T3–T4 interval than they were based on the data. This is due to the fact that all time spans are printed as equally broad, although the last time span comprised several years, being much broader than the other time intervals spanning less than half a year.

Latent Trait-Change Analyses

Throughout the next sections, we will present findings based on different analysis approaches. More specifically, we decided to provide preliminary analyses for the T1–T4 interval, followed by more detailed analyses across all measurement occasions. For all

analyses, we present both univariate and multivariate results. Within the manuscript, we mostly refer to results based on FIML to handle missing data. Results based on listwise deletion are presented in the supplementary document at the OSF (https://osf.io/pmy57/). Please note that all reported coefficients represent probabilistic point estimates, and can thus slightly vary across approaches.

Analyses across T1–T4. For preliminarily testing possible long-term sojourn effects, we ran univariate and multivariate latent trait-change model analyses across T1 and T4 only. Both the univariate and the multivariate latent models showed good fits to the data (see supplementary Table S2). While the univariate analyses yielded no meaningful sojourn effect across these two time points, the multivariate analysis suggested a long-term effect on openness change (see Table 3). This divergence is much likely attributable to low accuracy of the parameter estimates as indicated by broad 95% confidence intervals (CIs; Kelley & Rausch, 2006). Although 95% CIs were largely overlapping, only the one based on the multivariate analysis did not include zero: $b_{univariate}$ = .11 (95% CI [–.02, .25]), p = .103, vs. $b_{multivariate}$ = .15 (95% CI [.01, .29]), p = .031. Running the analyses with listwise deletion yielded the same pattern of results (see supplementary Table S3).

These initial analyses gave rise to the conclusion that there were no meaningful (lasting) sojourn effects on most if not all traits, except a small signal for openness. While these preliminary analyses provided a first overall impression on potentially missing or, if at all, marginal long-term effects of sojourns, it is also clear that only multiple time points with closer time intervals allow to disentangle the timing, magnitude, and trajectories at different stages during and after a sojourn in comparison to no sojourn (compare arguments by Luhmann et al., 2014, on the need of multiple time points when studying personality changes).

Analyses across T1–T2–T3–T4. All latent neighbor-change models across four time points fit well to the data (see supplementary Table S2). First, we report results on sojourn effects during and directly after a sojourn (i.e., T1–T2 and T2–T3). We then focus on effects between T3 and T4. Further sections are specifically dedicated to the role of age and sex differences as well as additional sojourns between T3 and T4.

Table 3
Sojourn Effects on Trait Changes as Derived from the Latent-Change Analyses Across T1–T4 Only

	Sojourn at T1						Sojourn between T3–T4					
Type of analysis	Univariate			Multivariate			Univariate			Multivariate		
Personality trait	*b*	*p*	95% CI	*b*	*p*	95% CI	*b*	*p*	95% CI	*b*	*p*	95% CI
Openness	.11	.103	–.02, .25	**.15**	.031	.01, .29	.07	.181	–.03, .16	.06	.182	–.03, .16
Conscientiousness	.05	.407	–.07, .18	.03	.616	–.10, .16	.06	.188	–.03, .15	.06	.226	–.03, .14
Extraversion	.07	.234	–.05, .19	.06	.331	–.06, .18	–.05	.258	–.13, .04	–.04	.329	–.12, .04
Agreeableness	.05	.533	–.10, .19	.03	.720	–.12, .17	.06	.283	–.05, .16	.05	.308	–.05, .15
Neuroticism	–.08	.236	–.22, .06	–.06	.395	–.20, .08	.07	.158	–.03, .17	.07	.154	–.03, .16

Note. N = 1,095. b = unstandardized effect estimate. Sojourn at T1 = sojourn in the academic term 2009/10; Sojourn between T3–T4 = sojourns since 10/01/2010. Unstandardized estimates are based on full information maximum likelihood to handle missing values. The latent variable indicators were standardized before the analyses. We controlled for the effects of sojourns between T3 and T4, age, and sex on latent trait changes.
For model fit indices, see supplementary Table S2.
Significant model parameters ($p < .05$) are shown in boldface.

Sojourn effects between T1 and T2. Zimmermann and Neyer (2013) have previously reported overall effects of sojourning, with different trends in openness, agreeableness, and neuroticism amongst sojourners compared to stayers. The present analyses based on FIML largely corroborated this pattern. More specifically, in the univariate and multivariate analyses, we found small positive sojourn effects on openness and agreeableness changes, and a negative sojourn effect on neuroticism change (see Table 4 for unstandardized sojourn effects). When running the same analyses with listwise deletion, four of the six small effects disappeared. That is, only the sojourn effect on openness change in the multivariate approach and the effect on neuroticism change in the univariate approach still reached statistical significance (see supplementary Table S4). As point estimates were quite comparable across FIML and listwise deletion, this divergence in statistical significance was most likely due to the lower statistical power of the smaller listwise-deletion sample. Accordingly, the confidence intervals revealed higher accuracy in the FIML-based analyses.

Figure 4 is based on the latent modeling results of the univariate and multivariate analyses. It suggests an increase in sojourners' and a decrease or no change in stayers' openness as well as an increase in sojourners' and a decrease in stayers' agreeableness across T1–T2. Like Figure 3, it shows accelerated neuroticism decrease in sojourners, with no substantial change in stayers. Please note that Figure 3 and Figure 4 are not directly comparable. While Figure 3 pictures the standardized mean levels based on the raw data, the trends in Figure 4 are based on the latent variable estimates controlled for measurement error and multiple covariates. In other words, Figure 4 shows hypothetical trends for an average sojourner and stayer.

To compare short- and long-term sojourners, we also tested for differences in their T1–T2 trait changes (three-group analyses). While most results did not indicate differences between the sojourn groups, the univariate analysis revealed a small positive sojourn effect with regard to short-term sojourners' extraversion change, with no respective effect for long-term sojourners (see supplementary Table S5 and Figure S3). However, this effect was not found in the multivariate approach.

Sojourn effects between T2 and T3. One important extension to the modeling approach used by Zimmermann and Neyer (2013) was the inclusion of an additional change variable that captured change between T2 and T3 in all participant groups. When only comparing sojourners and stayers, we did not find any effects of a sojourn on trait changes

Table 4
Sojourn Effects on Trait Changes as Derived from the Latent-Change Analyses Across Four Time Points

Type of analysis	Openness						Conscientiousness						Extraversion		
	Univariate			Multivariate			Univariate			Multivariate			Univariate		
Effects	*b*	*p*	95% CI	*b*	*p*	95% CI	*b*	*p*	95% CI	*b*	*p*	95% CI	*b*	*p*	95% CI
Sojourn at T1 →															
ΔT1–T2	**.09**	.002	.03, .16	**.10**	.002	.04, .16	.05	.125	–.01, .11	.04	.200	–.02, .10	.02	.544	–.04, .08
ΔT2–T3	–.02	.521	–.08, .04	–.01	.642	–.07, .05	.01	.841	–.06, .07	.00	.907	–.06, .07	–.01	.610	–.07, .04
ΔT3–T4	.03	.624	–.10, .16	.07	.319	–.07, .20	.01	.903	–.12, .13	.00	.962	–.12, .13	.08	.185	–.04, .20
Sojourn between T3–T4 →															
ΔT3–T4	.08	.079	–.01, .17	.07	.101	–.01, .16	.05	.271	–.04, .13	.05	.291	–.04, .13	–.03	.401	–.11, .05

Note. N = 1,095. *b* = unstandardized effect estimate. Sojourn at T1 = sojourn in the academic term 2009/10; Sojourn between T3–T4 = sojourns since 10/01/2010.
Unstandardized estimates are based on full information maximum likelihood to handle missing values. The latent variable indicators were standardized before the analyses.
We controlled for the effects of sojourns between T3 and T4, age, and sex on latent trait changes.
For model fit indices, see supplementary Table S2.
Significant model parameters ($p < .05$) are shown in boldface.

Table 4 *continued*

Multivariate			Agreeableness						Neuroticism					
			Univariate			Multivariate			Univariate			Multivariate		
b	*p*	95% CI	*b*	*p*	95% CI	*b*	*p*	95% CI	*b*	*p*	95% CI	*b*	*p*	95% CI
.01	.788	–.05, .07	**.12**	.001	.05, .20	**.10**	.010	.02, .17	**–.13**	.001	–.20, –.06	**–.10**	.004	–.17, –.03
–.01	.644	–.07, .04	.02	.681	–.06, .09	.03	.478	–.05, .10	.01	.803	–.06, .07	–.01	.676	–.08, .05
.10	.115	–.02, .22	–.10	.153	–.23, .04	–.09	.226	–.22, .05	.04	.606	–.10, .17	.02	.728	–.11, .16
–.03	.441	–.11, .05	.08	.069	–.01, .17	.07	.116	–.02, .16	.07	.141	–.02, .15	.05	.246	–.04, .14

during this time span (see Table 4). To investigate the timing and trajectories of sojourners' trait changes more deeply, we also tested for differences between sojourn groups for the T2–T3 interval. The reason was twofold: First, during this period of time, a reversed effect of a short-term sojourn (one semester) on personality trait changes could indicate effects of the transition back home on trait changes for short-term sojourners. Second, the assessment of effects of a long-term sojourn on personality trait changes between T2 and T3 allowed us to investigate the timing or discontinuity of trait changes during a stay abroad. While the univariate model estimates revealed a relative increase in short-term sojourners'

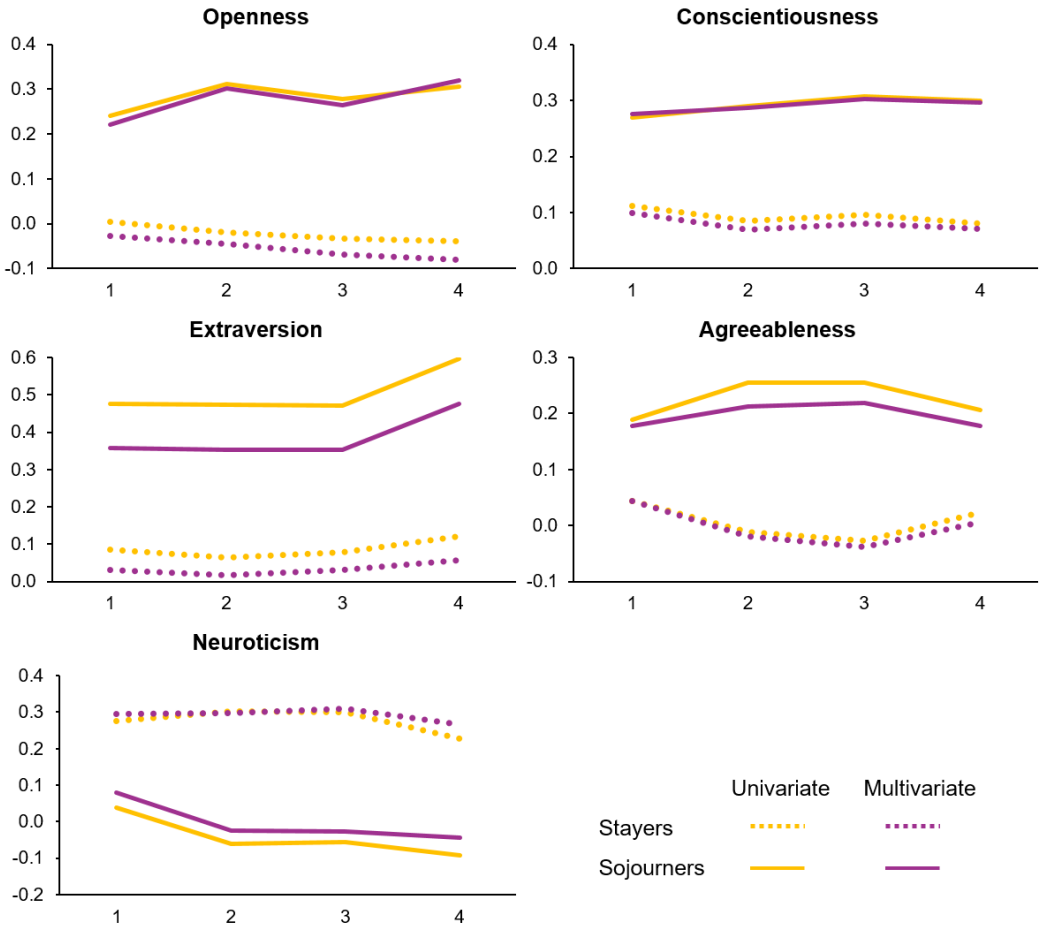

Figure 4. Big Five latent change trends of stayers and sojourners. We printed the respective estimates derived from the univariate and multivariate latent neighbor-change models after standardizing the latent factor indicators, centered on age and sex. T1 values are the unweighted means of the estimates that derived from additional regressions of the Big Five traits on sojourn status in our models, also centered on age and sex. This way, the line of stayers for example represents an imaginary person of age 22.59 years at study begin that did not live abroad. The x axis denotes time points. Please note that all time spans are printed as equally broad, although the last time span comprised several years, being much broader than the other time spans.

neuroticism after their return back home compared to the other groups (see supplementary Table S5 and Figure S3), the multivariate approach did not.

In sum, the analyses indicated that sojourn effects on changes primarily took place between T1 and T2 – that is, within the first five months of a sojourn – independently of the intended sojourn duration (short- or long-term sojourn). For most traits, spending additional months abroad or returning home did not add to or reverse the observed sojourn effects, indicating neither accentuation nor reversibility of effects (beyond a weak signal for returnees' neuroticism that relatively increased directly after their transition back home compared to the other groups' neuroticism).

Sojourn effects until T4? Significance levels of effects on trait changes T3–T4 in Table 4 suggested no differences in trait change trajectories between sojourners' and stayers' openness, conscientiousness, and neuroticism. This was also true for the analyses based on listwise deletion (supplementary Table S4). Although the point estimates for a sojourn effect on openness change between T1 and T2 ($b_{univariate}$ = .09, p = .002; $b_{multivariate}$ = .10, p = .002) shared a comparable magnitude with the effects on agreeableness change between T3 and T4 ($b_{univariate}$ = –.10, p = .153; $b_{multivariate}$ = –.09, p = .226), only the former effects statistically differed from zero (see Table 4). A similar picture emerged for the three-group analyses. Estimates of the sojourn effect on neuroticism change T3–T4 suggested a negative trend in short-term sojourners ($b_{univariate}$ = –.11, p = .233; $b_{multivariate}$ = –.10, p = .244) and a positive trend in long-term sojourners ($b_{univariate}$ = .14, p = .078; $b_{multivariate}$ = .12, p = .138; see supplementary Table S5 and Figure S3), suggesting more decrease in short-term sojourners' neuroticism, and a reversed effect for long-term sojourners after their return. However, none of the trends reached statistical significance, indicating that the T3–T4 effects might have failed to detect small effects due to power limitations.

We were not aware of any procedure to reliably estimate the power of our latent model. As an approximation, we applied a power analysis using G*Power 3.1.9.4 (Faul, Erdfelder, Buchner, & Lang, 2009) based on the manifest effect sizes. This analysis revealed a substantial lack of sensitivity regarding sojourn effects on personality trait changes during the T3–T4 time span. That is, even if there were true effects, they might not have reached statistical significance because of the smaller sample size. Of course, we cannot directly translate the findings of a power analysis of manifest effects to a latent modeling approach, and the FIML-based latent variable analysis might have been more sensitive as the smaller

95% CIs suggested. For example, we received significant results for T1–T2, although power for this time span was also estimated as being poor. However, we acknowledge that results of the manifest power analysis are a hint on reduced power for T3–T4 as compared to T1–T2. Accordingly, there were much wider 95% CIs for T3–T4 estimates compared to the other time spans (see Table 4). Therefore, effects across T3–T4 could not be estimated with the same precision as effects across the other time intervals (i.e., T1–T2, T2–T3; compare Kelley & Rausch, 2006). Please note that this also applies for the T1–T4 analyses (compare 95% CIs in Table 3). In view of these limitations, an interpretation of the effects across T1–T4 and T3–T4 that is solely based on statistical significance does not seem warranted.

Model-comparison analyses. In face of the aforementioned accuracy and power reasons, we pursued an alternative strategy to directly test for reversed, accentuated, and sustained effects using model comparison criteria. To that end, we ran formal model comparison analyses with and without constraining sojourn effects on trait change in the T3–T4 interval. Please note that we used the exact point estimates as they had been estimated previously by the univariate and multivariate approaches. Thus, point estimates could slightly differ between model comparisons in the two approaches.

We started by investigating whether effects on changes between T1 and T2 later reversed (*reversed effects*). For example, we tested whether equating the sojourn effect on openness change between T3 and T4 to its receptive reverse value of the T1–T2 interval (i.e., $b_{univariate}$ = –.09 and $b_{multivariate}$ = –.10; see Table 4) led to a significant decrease in model fit. Second, we tested whether effects on T1–T2 changes might have occurred again between T3 and T4 with the same size (*accentuated effects*).[10] For example, we tested whether equating the sojourn effect on openness change T3–T4 to the effect on openness change T1–T2 (i.e., $b_{univariate}$ = .09 and $b_{multivariate}$ = .10) led to a significant decrease in model fit. Third, we tested whether we could fix sojourn effects on changes between T3 and T4 to zero (*sustained effects*). As we had not identified significant effects on changes between T2 and T3, [11] we

[10] We are aware that the assumption of same-size effects for the substantially longer time interval between T3 and T4 is somewhat arbitrary as little is known about the exact shape and timing of change in the different traits over the course of young adulthood. However, previous findings supported the assumption of unidirectional change in all traits between age 20 and 30 (Specht et al., 2011). Hence, assuming the same amount of change for the 6-months interval from T1-T2 and the four years between T3-T4 was deemed the most conservative prognosis.

[11] Although there was a positive effect on neuroticism change in returned short-term sojourners in the univariate model, there were no sojourn effects for the whole group of sojourners during this time span.

equated all effects across T2–T3 to zero for these additional analyses (the same strategy was applied for the three-group models).

Reversed effects under investigation. Using five separate Wald tests of parameter constraints for both the univariate and the multivariate analyses, we tested whether effects of a sojourn on trait changes later reversed. For openness, we found a significant decrease in model fit in the multivariate analysis when constraining the data to the reversed effect, indicating that a reversed sojourn effect was unlikely (see Table 5). Although the same trend of model fit decline could be observed for the univariate model analysis, the decrease in model fit was not statistically significant. There were no further hints on model deteriorations regarding reversed effects in the other traits. Summed up, reversed sojourn effects tended to be unlikely for openness, but not for the other traits.

Accentuated effects under investigation. Wald tests of parameter constraints were significant for agreeableness in both the univariate and multivariate model analysis approaches, indicating that an accentuated-effects model did not fit this trait's pattern well. In other words, an accentuated sojourn effect on agreeableness change between T3 and T4 was statistically not likely. Moreover, the Wald test was significant for neuroticism in the univariate analysis, but not in the multivariate analysis. To sum up, model-comparison analyses indicated that an accentuated sojourn effect on agreeableness was unlikely, while it only tended to be unlikely for neuroticism.

Sustained effects under investigation. We repeated the model tests for sustained effects, but did not find any evidence for a substantial decrease in model fit (see Table 5). Thus, we cannot rule out sustained effects for any trait. Comparing each trait regarding the three effect patterns (when freeing effects for the other traits in the multivariate model) might yield a more differentiated picture.

Comparing reversed, accentuated, and sustained effects. The Bayesian Information Criterion (BIC) and the sample-size adjusted BIC (SABIC) were used to compare two independent (i.e., not nested) models with the same set of dependent variables. Although there is no cutoff criterion for either of them, a smaller BIC or SABIC value indicates a comparably better model fit (Kelloway, 2015). For the BIC, Lubke et al. (2017) have proposed $|\Delta BIC| \geq 2.00$ between two models as a rule of thumb to assume evidence against the model with the higher value (based on prior work by Kass & Raftery, 1995). Although we are not aware of any rule of thumb regarding

Table 5
Results of Investigating Long-Term Sojourn Effects More Deeply

	Openness					Conscientiousness					Extraversion				
	Wald tests			Bayesian criteria		Wald tests			Bayesian criteria		Wald tests			Bayesian criteria	
Statistics	$\Delta\chi^2$	*df*	*p*	BIC	SABIC	$\Delta\chi^2$	*df*	*p*	BIC	SABIC	$\Delta\chi^2$	*df*	*p*	BIC	SABIC
Univariate															
Rev.	3.04	1	.081	14351.06	14255.78	1.00	1	.318	15214.11	15118.82	2.53	1	.112	13350.63	13255.34
Accent.	1.14	1	.286	14349.17	14253.88	0.41	1	.522	15213.53	15118.24	0.86	1	.354	13348.96	13253.67
Sust.	0.12	1	.735	14348.15*	14252.86*	0.03	1	.858	15213.15	15117.86	1.59	1	.208	13349.68	13254.40
Multivariate															
Rev.	**5.92**	1	.015	71216.16	70193.41	0.50	1	.482	71210.77	70188.03	2.84	1	.092	71213.11	70190.36
Accent.	0.38	1	.539	71210.65*	70187.91*	0.34	1	.559	71210.62	70187.87	1.84	1	.175	71212.11	70189.36
Sust.	0.83	1	.364	71211.10*	70188.36*	0.00	1	.952	71210.28	70187.53	2.31	1	.129	71212.58	70189.84

Note. N = 1,095. Rev. = *Reversed effects*: Setting the effect between T3–T4 at the opposite of the respective effect between T1–T2; Accent. = *Accentuated effects:* Setting the effect between T3–T4 at the same as the respective effect between T1–T2; Sust. = *Sustained effects:* Setting the effect for T3–T4 to 0.
For model fit indices, see supplementary Table S6.
The number of * indicates the number of models that the indexed model is superior to, based on ΔBIC $\geq$ 2 and ΔSABIC $\geq$ 2, respectively.
Significant model parameters ($p < .05$) are shown in boldface.

Table 5 *continued*

Agreeableness					Neuroticism				
Wald tests			Bayesian criteria		Wald tests			Bayesian criteria	
$\Delta\chi^2$	*df*	*p*	BIC	SABIC	$\Delta\chi^2$	*df*	*p*	BIC	SABIC
0.18	1	.674	16863.86*	16768.58*	1.88	1	.170	15636.90*	15541.61*
10.02	1	.002	16873.60	16778.31	**6.76**	1	.009	15641.73	15546.45
1.88	1	.170	16865.57*	16770.28*	0.38	1	.539	15635.40*	15540.11*
0.15	1	.695	71210.43*	70187.68*	1.60	1	.206	71211.87	70189.13
6.45	1	.011	71216.68	70193.94	3.08	1	.079	71213.35	70190.60
1.15	1	.283	71211.43*	70188.68*	0.06	1	.806	71210.34*	70187.59*

the SABIC, this criterion penalizes sample size less (Kelloway, 2015), but is comparable to the BIC apart from that. For the present study, we therefore applied the rule of thumb for both the BIC and the SABIC. As both criteria led to the same results (with only slight differences in decimal places), we will only report the results based on ΔBIC here (see Table 5). Please note that the BIC tends to prefer models with less parameters, especially in small samples, and that fit indices are probabilistic, and not absolute, criteria (Lubke et al., 2017).

For openness, an accentuated long-term effect fit comparably better than a reversed effect in the multivariate ($\Delta BIC_{\text{reversed-accentuated, multivariate}} = 5.51$), but not in the univariate model ($\Delta BIC_{\text{reversed-accentuated, univariate}} = 1.89$). Comparing the sustained-effect model to the accentuated-effect model yielded no substantial difference between these models for openness ($\Delta BIC_{\text{accentuated-sustained, univariate/multivariate}} = 1.02/–0.45$), while comparing it to the reversed-effect model indicated a better fit for the sustained-effect model ($\Delta BIC_{\text{reversed-sustained, univariate/multivariate}} = 2.91/5.06$). Hence, the model comparisons corroborate a sustained or accentuated pattern, while we can most likely rule out a reversed sojourn effect on openness change.

For conscientiousness, there were no substantial differences between the fit of the reversed- or accentuated-effect model ($\Delta BIC_{\text{reversed-accentuated, univariate/multivariate}} = 0.58/0.15$) and the sustained-effect model ($\Delta BIC_{\text{reversed-sustained, univariate/multivariate}} = 0.96/0.49$, and $\Delta BIC_{\text{accentuated-sustained, univariate/multivariate}} = 0.38/0.34$). For extraversion ($\Delta BIC_{\text{reversed-accentuated, univariate/multivariate}} = 1.67/1.00$, $\Delta BIC_{\text{reversed-sustained, univariate/multivariate}} = 0.95/0.53$, and $\Delta BIC_{\text{accentuated-sustained, univariate/multivariate}} = –0.72/–0.47$), the sustained model did not differ markedly from the other models, either. This indicated no substantial differences between all three patterns for conscientiousness and extraversion.

For agreeableness, a reversed-effect model ($\Delta BIC_{\text{reversed-accentuated, univariate/multivariate}} = –9.74/–6.25$) and a sustained-effect model ($\Delta BIC_{\text{accentuated-sustained, univariate/multivariate}} = 8.03/5.25$) both fit better than an accentuated-effect model. There was no substantial difference between the reversed-effect and the sustained-effect model ($\Delta BIC_{\text{reversed-sustained, univariate/multivariate}} = –1.71/–1.00$). For neuroticism, the sustained-effect model ($\Delta BIC_{\text{accentuated-sustained, univariate/multivariate}} = 6.33/3.01$) fit better than the accentuated-effect model. While the results were indifferent for comparing the reversed- and the accentuated-effect model ($\Delta BIC_{\text{reversed-accentuated, univariate/multivariate}} = –4.83/–1.48$), reversed- and sustained-effect models did not differ markedly ($\Delta BIC_{\text{reversed-sustained, univariate/multivariate}} = 1.50/1.53$). That is, reversed or sustained patterns seem

most adequate, while we can most likely rule out accentuated effects on agreeableness and neuroticism changes.

To conclude, these Bayesian model fit comparisons yielded evidence against reversed sojourn effects on openness change, as well as against accentuated effects on agreeableness and neuroticism changes. Findings for conscientiousness and extraversion were inconclusive. Yet, in the absence of initial sojourn effects for these traits, comparisons were somewhat arbitrary.

Summary of sojourn effects. After small initial sojourn effects on openness, agreeableness, and neuroticism changes between T1 and T2, there were no significant effects on further trait changes across T2–T3 or T3–T4 in any of the traits investigated (Table 4). Figure 4 shows the trait change patterns of stayers and sojourners based on the latent neighbor-change model analyses. For openness, the latent change models revealed substantial positive sojourn effects, with slightly differing estimates for T3–T4 across the univariate (b = .03) and the multivariate (b = .07) analyses. Wald tests as well as Bayesian model comparisons suggested that a reversed effect was unlikely, while an accentuated or sustained pattern were better suited to describe the data for openness. For neuroticism, the (insignificant) coefficients for T3–T4 pointed in the opposite direction ($b_{\text{univariate}}$ = .04, $b_{\text{multivariate}}$ = .02) than for T1–T2 ($b_{\text{univariate}}$ = –.13, $b_{\text{multivariate}}$ = –.10). Yet, these coefficients were substantially smaller than the T1–T2 effects and might thus be negligible. Wald tests along with Bayesian model comparisons suggested that an accentuated effect was less likely than a sustained or reversed pattern.

In contrast to stayers, sojourners revealed an increase in agreeableness between T1 and T2 ($b_{\text{univariate}}$ = .12, $b_{\text{multivariate}}$ = .10). The opposite trend ($b_{\text{univariate}}$ = –.10, $b_{\text{multivariate}}$ = –.09) appeared after T3 (although coefficients did not statistically differ from zero): Stayers, but not sojourners, showed a slight increase in agreeableness (see Figure 4). The confidence intervals in Table 4 (95% $CI_{\text{univariate}}$ [–.23; .04], 95% $CI_{\text{multivariate}}$ [–.22; .05]), accompanied by significant Wald tests of parameter constraints and Bayesian model comparisons indicated that an accentuated effect on agreeableness change was unlikely, but a sustained or reversed pattern better fit the data.

For conscientiousness and extraversion, all effects from the univariate and the multivariate models were close to zero. Accordingly, the model comparisons did not favor

any of the three potential patterns for these traits. Hence, conscientiousness and extraversion changes seem to be unaffected by sojourn experiences in the long run.

The role of sojourns between T3 and T4. Besides the aforementioned effects, a previous sojourn during the academic term 2009/10 had a positive effect on the number of future stays abroad between T3 and T4 [b = .36, p <.001]. Although the 95% CIs slightly differed, this effect was independent of the analysis strategy. Further, point estimates suggested that sojourns between T3 and T4 tended to be associated with small differences in changes in openness, agreeableness, and neuroticism (see Table 4). Although differences between those who moved abroad between T3 and T4 and those who did not were not significant, 95% CIs suggested that these effects might be in a positive range for openness ($b_{univariate}$ = .08 [–.01, .17], $b_{multivariate}$ = .07 [–.01, .16]), agreeableness ($b_{univariate}$ = .08 [–.01, .17], $b_{multivariate}$ = .07 [–.02, .16]), and neuroticism ($b_{univariate}$ = .07 [–.02, .15], $b_{multivariate}$ = .05 [–.04, .14]). This implies a tendency for accentuated or at least sustained differences in openness and agreeableness, but reversed effects for neuroticism in association with further sojourns.

Age and sex differences. In our analyses, we allowed for age and sex effects on all latent Big Five trait change variables. Tables 6 and 7 contain the latent regression effects of these covariates on trait changes for both the univariate and multivariate analyses across T1–T4 and T1–T2–T3–T4, respectively.

Sex differences. Sex differences were found for openness and neuroticism. In the T1–T4 approach, the multivariate analysis revealed that females tended to decrease more strongly or increased less strongly in openness than males, while the effect was not significant in the univariate analysis (see Table 6). In the T1–T2–T3–T4 approach, women's openness was found to decrease less or increased more than men's openness across T2–T3 in both the univariate and multivariate analyses (see Table 7). The multivariate analysis additionally suggested that females' openness decreased more or increased less than males' openness during T1–T2.

Across the T1–T4 interval, the sex effect on neuroticism change was the most pronounced effect (see Table 6), indicating that women decreased less steeply or increased more steeply in this trait than men. The pattern that females' neuroticism decreased less or increased more than that of their male counterparts was corroborated by positive sex effects for T1–T2 in the T1–T2–T3–T4 approach (see Table 7). We also found a positive sex effect

Table 6

Age and Sex Effects on Trait Changes as Derived from the Latent-Change Analyses Across T1–T4 Only

	Openness			Conscientiousness			Extraversion			Agreeableness			Neuroticism		
Effect	*b*	*p*	95% CI	*b*	*p*	95% CI	*b*	*p*	95% CI	*b*	*p*	95% CI	*b*	*p*	95% CI
Univariate															
Sex →	–.13	.105	–.28, .03	.04	.624	–.11, .18	.05	.432	–.08, .19	–.00	.979	–.16, .16	**.30**	<.001	.14, .46
Age →	.01	.479	–.02, .04	–.02	.087	–.05, .00	–.01	.514	–.03, .02	–.01	.721	–.03, .02	.01	.300	–.01, .04
Multivariate															
Sex →	**–.17**	.035	–.34, –.01	.02	.789	–.13, .17	.03	.714	–.12, .17	–.01	.924	–.18, .16	**.33**	<.001	.17, .50
Age →	.01	.506	–.02, .03	–.02	.145	–.04, .01	–.01	.569	–.03, .02	–.01	.668	–.03, .02	.01	.308	–.01, .04

Note. N = 1,095. b = unstandardized effect estimate. Unstandardized estimates are based on full information maximum likelihood to handle missing values. The latent variable indicators were standardized before the analyses. Sex was coded 0 (*male*), 1 (*female*). Age was centered before the analysis.

For model fit indices, see supplementary Table S2.

Significant model parameters ($p < .05$) are shown in boldface.

Table 7
Age and Sex Effects on Trait Changes as Derived from the Latent-Change Analyses Across Four Time Points

Type of analysis	Openness						Conscientiousness						Extraversion		
	Univariate			Multivariate			Univariate			Multivariate			Univariate		
Effects	*b*	*p*	95% CI	*b*	*p*	95% CI	*b*	*p*	95% CI	*b*	*p*	95% CI	*b*	*p*	95% CI
Sex →															
ΔT1–T2	–.07	.061	–.14, .00	**–.10**	.016	–.17, –.02	.02	.568	–.05, .10	.04	.307	–.04, .12	.05	.212	–.03, .12
ΔT2–T3	**.08**	.027	.01, .15	**.10**	.005	.03, .18	–.05	.171	–.13, .02	–.05	.225	–.12, .03	–.03	.448	–.09, .04
ΔT3–T4	–.10	.172	–.25, .04	–.14	.088	–.29, .02	.07	.297	–.07, .21	.07	.335	–.04, .22	.05	.499	–.09, .18
Age →															
ΔT1–T2	.00	.587	–.01, .02	.00	.463	–.01, .02	–.00	.737	–.01, .01	–.00	.883	–.01, .01	.00	.938	–.01, .01
ΔT2–T3	–.01	.271	–.02, .01	–.01	.292	–.02, .01	.00	.568	–.01, .02	.00	.462	–.01, .02	–.01	.120	–.02, .00
ΔT3–T4	.02	.128	–.01, .04	.02	.133	–.01, .04	**–.03**	.027	–.05, –.00	**–.03**	.041	–.05, –.00	.01	.648	–.02, .03

Note. N = 1,095. b = unstandardized effect estimate. Unstandardized estimates are based on full information maximum likelihood to handle missing values. The latent variable indicators were standardized before the analyses. Sex was coded 0 (*male*), 1 (*female*). Age was centered before the analysis.
For model fit indices, see supplementary Table S2.
Significant model parameters ($p < .05$) are shown in boldface.

Table 7 *continued*

Multivariate			Agreeableness						Neuroticism					
			Univariate			Multivariate			Univariate			Multivariate		
b	*p*	95% CI	*b*	*p*	95% CI	*b*	*p*	95% CI	*b*	*p*	95% CI	*b*	*p*	95% CI
.04	.305	–.04, .11	.00	.985	–.09, .09	.07	.158	–.03, .16	**.11**	.017	.02, .19	**.16**	<.001	.07, .25
–.03	.392	–.10, .04	.03	.467	–.05, .12	.03	.558	–.06, .12	.01	.963	–.08, .08	–.02	.631	–.10, .06
.04	.553	–.10, .18	.02	.786	–.13, .17	.02	.858	–.15, .18	**.19**	.017	.03, .35	.14	.083	–.02, .30
.00	.594	–.01, .01	–.00	.597	–.02, .01	–.00	.605	–.02, .01	.01	.238	–.01, .02	.01	.416	–.01, .02
–.01	.135	–.02, .00	–.00	.813	–.02, .01	–.00	.791	–.02, .01	–.00	.856	–.01, .01	.00	.971	–.01, .01
.01	.601	–.02, .03	–.01	.569	–.03, .02	–.01	.533	–.03, .02	.01	.384	–.01, .04	.01	.289	–.01, .04

for T3–T4 in the univariate analysis with a similar trend in the multivariate analysis (see Table 7), indicating less decrease or more increase in women's neuroticism.

Age differences. Age was not found to predict any trait changes across T1–T4. The T1–T2–T3–T4 analyses revealed a small negative age effect on conscientiousness change during T3–T4. Compared to a slight increase for the average person (see Table 2), this indicated that being older was associated with less increase in conscientiousness, possibly even indicating a decrease for older people.

Discussion

Do sojourn effects on personality trait changes last? The aim of our study was to examine long-term trait changes of sojourners in comparison to their fellow students who did not live abroad. Our results indicated that studying abroad might promote differences in openness, agreeableness, as well as neuroticism changes during the first few months of a sojourn. Differences between the sojourn groups in their T1–T2 extraversion change and their T2–T3 neuroticism change only occurred in the univariate models. After that, there were no substantial group differences in trait changes across five years. However, our study had to deal with lower power and accuracy to detect effects across the follow-up time span between T3 and T4. At the descriptive level, point estimates of sojourn effects derived from the latent change analyses suggested a sustained or accentuated pattern for openness, a sustained effect on neuroticism change, and a reversed effect on agreeableness change. In addition, the 95% CIs and Bayesian model comparisons suggested that a reversed effect was unlikely for openness, and accentuated sojourn effects were unlikely for agreeableness and neuroticism. In the following, we will discuss our findings in more detail.

Sojourn Effects Between T1 and T2

We showed that sojourn effects on openness, agreeableness, and neuroticism changes occurred early within the first five months of sojourn-related context changes and experiences, independent of the intended sojourn duration (one or two semesters abroad). However, the effects were small. The fact that small sojourn effects were observed during the first five months indicated that trait changes can occur relatively quickly, in contrast to what personality psychologists have previously proposed for the effects of life experiences (see also Roberts, Luo et al., 2017). Ward et al. (1998) reported highest adaptation difficulties in the beginning of a sojourn with steep decreases thereafter. Although adaptation difficulties are not exactly translatable to personality traits, this might be a hint on early adaptation processes

that have led to personality trait changes between T1 and T2 (compare Roberts, 2018; Roberts & Jackson, 2008).

It is important to stress that this study was not able to fully capture the dynamics of student sojourns on trait changes with regard to potential anticipation effects. For example, it is possible that sojourners experienced more anxiety prior to departure than they usually did (see Suanet & van de Vijver, 2009, on the adaptation to a new social context). If that was the case, sojourn effects during the first five months abroad would rather reflect returns to the baseline instead of maturation patterns. However, there are several arguments that challenge these speculations: First, while it may seem plausible to assume increased levels of anxiety (and thus respective state changes in neuroticism) prior to departure, the patterns for other traits are less clear. For example, it is less evident why levels of openness and agreeableness should decrease in advance of a stay abroad.

Furthermore, some authors have argued that mood disturbances might last for 4 to 6 months during an adaptation to a new environment abroad (see Ward et al., 1998). If changes in neuroticism merely reflected state changes in negative mood, it would be more plausible to expect a positive effect of sojourning on neuroticism change (i.e., an increase). To more thoroughly investigate these issues, longitudinal studies that cover the time before the departure (e.g., by implementing a waiting-group design) are needed.

Effects of the Transition Back Home?

We did not find any support for sojourn effects between T2 and T3 for long-term sojourners. However, the analyses revealed some hints but an inconsistent signal for a positive sojourn effect on short-term sojourners' neuroticism change T2–T3, indicating possible adaptation problems for returnees. This is in line with theories on negative effects of the transition back home on trait changes, previously described as "reverse culture shock" (Christofi & Thompson, 2007, p. 53).

Long-Term Personality Trait Changes of Sojourners and Stayers

The main focus of the present study was to investigate whether sojourn effects on personality trait changes might last. This is interesting both from a theoretical and an empirical perspective. For employers in the German economy, for example, it is of as much (or even more) interest how sojourners develop in the long run (DAAD, 2015). Likewise, previous studies identified a lack of research concerning the long-term effects of life

experiences in the scientific literature (Bleidorn et al., 2018, 2020). We tried to fill this gap with respect to sojourn experiences. However, we cannot fully disentangle the three potential patterns of reversed, accentuated, and sustained sojourn effects based on the current follow-up sample at hand. Nevertheless, we believe that we can discuss more or less likely patterns based on point estimates derived from the multivariate latent change model, their 95% CIs, and additional (Bayesian) model comparisons.

Regarding openness, we interpret our findings as most probably reflecting no reversed differences across the groups. This finding provides some support for sustained or accentuated sojourn effects on openness, above and beyond possible effects of additional sojourns. It is consistent with effects of clinical interventions on personality trait changes that were shown to mainly occur within the first weeks and to persist later on (Roberts, Luo et al., 2017). However, the obtained results do not allow for conclusive answers as regards the interpretation of the effect as a sustained or accentuating pattern.

The visualized pattern suggested a reversed sojourn effect on agreeableness change that was attributable to a possible slight increase in stayers across T3 and T4, with no change or a slight decrease in sojourners across this period (compare Figure 4). This pattern was corroborated by inspection of confidence intervals and model comparisons that indicated no accentuated effect on agreeableness change. However, we cannot fully rule out sustained differences between sojourners and stayers regarding their agreeableness changes.

We found no effect on neuroticism change between T3 and T4. Although manifest mean-level trends suggested a tendency toward a reversed sojourn effect, the point estimate based on the latent neighbor-change models did not clearly support this suggestion as the coefficient was close to zero (see Table 4, and compare Figure 3 to Figure 4). Please note that we cannot rule out a reversed effect especially for the long-term sojourners, but we can most likely rule out an accentuated long-term effect on neuroticism change. This pattern might be interpreted as an accelerated maturation between T1 and T2 towards more emotional stability among sojourners compared to stayers, with short-term neuroticism increases for returnees, and stayers – if at all – slowly and incrementally catching up since T3, but still not having reached sojourners' mean level by the end of the study.

The advantage of latent modeling. Against the backdrop of the manifest mean-level trends, we acknowledge the possibility that openness, agreeableness, and neuroticism all showed reversed sojourn effects. However, based on the latent variable analyses, which

provided more precise estimates controlled for error of measurement, several covariates, and inter-trait correlations, we can most probably rule out a reversed effect of sojourning on openness (and accentuated effects of sojourning on agreeableness and neuroticism changes) over a time span of five years. In sum, traits seem to follow different change patterns. This trait dependence of life-event effects has repeatedly been reported (Bleidorn et al., 2018; Specht et al., 2011). However, we still need to stress that our findings can only be seen as trends with limited statistical back-up, and that they need to be replicated by future research.

The role of recurring sojourns. The present study allowed us to replicate former studies regarding the finding that earlier sojourns are predictors of future sojourns (Netz & Jaksztat, 2014; Niehoff et al., 2017). Point estimates (see Table 4) also revealed that sojourns between T3 and T4 might have positive effects on trait changes between T3 and T4 in openness, agreeableness, and neuroticism after the control of trait levels at T3. For openness and agreeableness, these effects were comparable in size to the significant effects found for the T1–T2 interval, but were not statistically significant due to reduced power and accuracy. Moreover, the findings are limited by the fact that we did not know much about the exact timing and circumstances of those stays abroad. The developmental interplay of repeated mobility experiences during different phases of the educational career and across the lifespan remains an interesting objective to be more thoroughly explored by future studies. For example, previous research showed that first-time sojourners benefit more from participation in international student mobility than experienced sojourners (Zimmermann, Greischel, & Jonkmann, 2020). Hence, it might be valuable to more elaborately examine if a similar pattern occurred with regard to development in the Big Five traits, thereby comparing individual patterns of sojourn effects on trait changes with regard to the timing of a (first) sojourn.

Limitations and Implications for Future Directions

Although our longitudinal design and analytical approach have several strengths, some limitations need to be addressed. First, a major restriction were the lower power and accuracy to detect effects for the T1–T4 and the T3–T4 intervals due to the reduced T4 sample size. We tried to accommodate this limitation with extensive additional analyses and model comparisons, and aspired to interpret our findings in a balanced and cautious way. Nevertheless, future studies should try to replicate our findings based on larger samples with more data points.

Second, one important limitation is our non-experimental design (see also Zimmermann & Neyer, 2013), whereby participants selected themselves into the groups of stayers and sojourners according to their personal characteristics. Although we included initial trait levels in our analysis, we cannot rule out that the observed effects resulted from influences that were connected to those trait levels, but that were not assessed in the current study design. Moreover, we do not exactly know if the observed effects are necessarily due to socialization effects. That is, it might be the case that differences between stayers and sojourners that had led to the observed trends were independent of the sojourn itself or had already led to the decision for a sojourn. However, both the distinctive pattern of socialization beyond initial trait levels as well as the accumulating amount of studies that corroborated the importance of sojourn experiences with regard to personality or identity development (Greischel et al., 2016; Greischel, Noack, & Neyer, 2018; Niehoff et al., 2017) support an interpretation of our results in terms of sojourn effects. In this regard, studies with waiting-group designs would be helpful as comparing trajectories of present and future sojourners (that are likely to be very similar with regard to all characteristics but the exact timing of their sojourn) may help to more thoroughly investigate the interplay of selection and socialization effects. One further alternative in this regard are large representative panel studies, in which more adequate control groups of sojourners can be identified with the use of propensity score matching. However, to our knowledge, such panel data are currently not available.

Third, our design did not allow us to gather information on the mechanisms that account for the observed pattern of socialization effects. Our findings suggested reduced openness and agreeableness decreases and accentuated decrease in neuroticism directly after sojourn-induced contextual changes. As a consequence, investigating sojourners' psychological and sociocultural adaptation (as an indicator of their successful mastery of the sojourn demands) might provide insights into the mediators of personality development. Studies on identity development in the context of high school students' sojourn experiences corroborated the importance of these mechanisms (Greischel et al., 2018; Greischel, Noack, & Neyer, 2019).

Moreover, from a micro-analytical perspective, the high-frequency examination of concrete behavior (changes) (e.g., with ambulatory assessment methods) might be a promising way to assess and integrate the short- and long-term processes of personality trait change (see Geukes, van Zalk, & Back, 2018; Wrzus, & Roberts, 2017). With regard to the micro-processes that occur in the context of (dyadic) interactions, the PERSOC model (Back

et al., 2011) provides an encompassing description of the social interaction units that relate to (changes in) self- and relationship dispositions, and thus account for personality changes.

In addition, females comprised almost 80% of the participants, heavily skewing the sample composition. To address this, we controlled for sex effects (as well as for age effects) in our analysis. As we found sex to be associated with changes in openness and neuroticism, we recommend exploring these differences in future research. In addition, the sample was limited to German students that mostly went to European countries for one or two semesters (see also Zimmermann & Neyer, 2013). We therefore cannot generalize our findings to students from other countries or students who stayed in other host countries. Further research has already shown that these variables are of high importance: For example, the cultural distance between one's home and host country seems to play a major role in the adjustment processes of sojourners (e.g., Suanet & van de Vijver, 2009). In Denissen et al.'s (2013) self-regulation theory, cultural values are an important predictor of personality changes. An examination of the difference between home and host cultural values could be a promising way to better understand the process of value acceptance in the new environment as a source of adaptation and trait changes.

From the self-regulation perspective, one might also argue that sojourners have experienced a shift in their reference values, i.e., by engaging in new behaviors and ideas, exploring (cultural) diversity, and handling the (social) challenges of living abroad (Denissen et al., 2013). In turn, discrepancies between these values and their observed behavior might have led sojourners to change their habits and daily behaviors throughout their sojourn. Now, several years later, some of these shifted values might still be in place and regulate their day-to-day behavior, while others do not. An investigation of (changes in) reference values and their contingency upon sojourners' goals that motivated their stay abroad (Zimmermann, Schubert, Bruder, & Hagemeyer, 2017) would help to clarify these processes.

Further, psychological measures that rely on self-ratings are often flawed, for example, by a social desirability bias. Recent research suggests that the use of both self- and informant ratings produces a more valid measure of personality traits, and better detects different perspectives on personality changes (Luan, Hutteman, Denissen, Asendorpf, & van Aken, 2017). As the present study only used self-report data on personality traits, the possibility of bias effects has to be taken into account when interpreting the findings.

Finally, there is evidence that differences *within* a sample could have an impact on trait changes as well. Denissen, Luhmann, Chung, and Bleidorn (2019) have recently reported significant variation in individuals' reactions to life events with regard to personality trait changes. In fact, Niehoff et al. (2017) have reported similar results on sojourn effects. Besides slight differences in the traits that were associated with selection and socialization effects, the authors found that the 25% of sojourners with the highest conscientiousness and agreeableness levels before their sojourn changed in these traits in the direction of more average trait levels. This illustrates that sojourn benefits in terms of personality maturation might depend upon pre-departure personality constellations. Our focus in the present research was on the exploration of developmental differences between sojourners and stayers. However, future research may follow up on these findings and more thoroughly explore conditions of differential development within the group of sojourners.

Conclusions

International mobility is a prevalent life event among university students in industrialized societies. Our study suggested that personality trait changes in association with student sojourns occur early and are small. With regard to the question if sojourn effects on personality change last, we could most likely rule out a reversed sojourn effect on openness, but considered the possibility of sustained or accentuated long-term differences. Likewise, we could most likely rule out accentuated differences between sojourners' and stayers' agreeableness changes over the course of five years. By contrast, a reversed pattern seemed most likely to describe the data. Although we could also most likely rule out an accentuated effect on differences between sojourners' and stayers' neuroticism changes, reversed and sustained differences appeared to be equally likely for the group of sojourners as a whole.

Separating short-term from long-term sojourners revealed possible effects of the type of sojourn. More specifically, while short-term sojourners' extraversion increased relatively to the other groups across T1–T2, returning home tended to inhibit neuroticism decrease for short-term sojourners. These findings were only found in the univariate analysis, but might be interesting starting points for future investigations. In addition, we found some hints that multiple sojourns might have additional effects on trait change over the course of five years, i.e., they might sustain or accentuate in openness and agreeableness change.

Our results can only be seen as first hints on long-term sojourn effects on personality trait changes, and may help future studies to generate hypotheses on the magnitude, stability,

and directions of differences between sojourners and stayers (Kenny & Judd, 2019). Future studies should replicate the trends found in this study and test our assumptions with more robust designs (e.g., waiting-group designs) and larger samples. An important objective for future investigations might also be a narrower examination of personality-environment transactions by taking further individual variables (e.g., identity or acculturation motivation) and environmental aspects (e.g., cultural differences between host countries, social background) into account (Wagner, Orth, Bleidorn, Hopwood, & Kandler, 2020). This seems promising in achieving a deeper understanding of the underlying adaptation processes that drive personality trait changes during and after sojourn experiences and, maybe, life experiences in general.

References

Back, M. D., Baumert, A., Denissen, J. J. A., Hartung, F. M., Penke, L., Schmukle, S. C.,. . . Wrzus, C. (2011). PERSOC: A unified framework for understanding the dynamic interplay of personality and social relationships. *European Journal of Personality*, 25, 90–107. https://doi.org/10.1002/pesr.811

Bleidorn, W., Hopwood, C. J., Back, M. D., Denissen, J. J. A., Hennecke, M., Jokela, M., . . . Zimmermann, J. (2020). Longitudinal Experience-Wide Association Studies (LEWAS) – A framework for studying personality change. *European Journal of Personality*. https://doi.org/10.1002/per.2247

Bleidorn, W., Hopwood, C. J., & Lucas, R. E. (2018). Life events and personality trait change. *Journal of Personality.* Advance online publication. https://doi.org/10.1111/jopy.12286

Bleidorn, W., Kandler, C., Riemann, R., Spinath, F. M., & Angleitner, A. (2009). Patterns and sources of adult personality development: Growth curve analyses of the NEO PI-R scales in a longitudinal twin study. *Journal of Personality and Social Psychology, 97*, 142–155. https://doi.org/10.1037/a0015434

Bleidorn, W., Klimstra, T. A., Denissen, J. J. A., Rentfrow, P. J., Potter, J., & Gosling, S. D. (2013). Personality maturation around the world: A cross-cultural examination of social-investment theory. *Psychological Science*, *24*, 2530–2540. https://doi.org/10.1177/0956797613498396

Christofi, V., & Thompson, C. L. (2007). You cannot go home again: A phenomenological investigation of returning to the sojourn country after studying abroad. *Journal of Counseling & Development, 85*, 53–63. https://doi.org/10.1002/j.1556-6678.2007.tb00444.x

DAAD (2015). DAAD-IW-Studie 2015 [DAAD-IW study 2015]. Retrieved from https://eu.daad.de/service/auswertung-und-statistik/studien-und-auswertungen-der-na-daad/daad-iw-studie-2015/de/49310-daad-iw-studie-01/

Denissen, J. J. A., Luhmann, M., Chung, J. M., & Bleidorn, W. (2019). Transactions between life events and personality traits across the adult lifespan. *Journal of Personality and Social Psychology*, *116*, 612–633. https://doi.org/10.1037/pspp0000196

Denissen, J. J. A., Ulferts, H., Lüdtke, O., Muck, P. M., & Gerstorf, D. (2014). Longitudinal transactions between personality and occupational roles: A large and heterogeneous study of job beginners, stayers, and changers. *Developmental Psychology*, *50*, 1931–1942. https://doi.org/10.1037/a0036994

Denissen, J. J. A., van Aken, M. A. G., Penke, L., & Wood, D. (2013). Self-regulation underlies temperament and personality: An integrative developmental framework. *Child Development Perspectives*, *7*, 255–260. https://doi.org/10.1111/cdep.12050

Enders, C. K. (2011). Analyzing longitudinal data with missing values. *Rehabilitation Psychology, 56*, 267–288. https://doi.org/10.1037/a0025579

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*, 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Geiser, C., & Lockhart, G. (2012). A comparison of four approaches to account for method effects in latent state–trait analyses. *Psychological Methods, 17*, 255–283. https://doi.org/10.1037/a0026977

George, L. G., Helson, R., & John, O. P. (2011). The "CEO" of women's work lives: How Big Five Conscientiousness, Extraversion, and Openness predict 50 years of work experiences in a changing sociocultural context. *Journal of Personality and Social Psychology*, *101*, 812–830. https://doi.org/10.1037/a0024290

Geukes, K., van Zalk, M., & Back, M. D. (2018). Understanding personality development: An integrative state process model. *International Journal of Behavioral Development*, *42*, 43–51. https://doi.org/10.1177/0165025416677847

Graham, E. K., Weston, S. J., Gerstorf, D., Yoneda, T. B., Booth, T., Beam, C. R., . . . Mroczek, D. K. (2020). Trajectories of big five personality traits: A coordinated analysis of 16 longitudinal samples. *European Journal of Personality*, https://doi.org/10.1002/per.2259

Greischel, H., Noack, P., & Neyer, F. J. (2016). Sailing uncharted waters: Adolescent personality development and social relationship experiences during a year abroad. *Journal of Youth and Adolescence*, *45*, 2307–2320. https://doi.org/10.1007/s10964-016-0479-1

Greischel, H., Noack, P., & Neyer, F. J. (2018). Oh, the places you'll go! How international mobility challenges identity development in adolescence. *Developmental Psychology*, *54*, 2152–2165. https://doi.org/10.1037/dev0000595

Greischel, H., Noack, P., & Neyer, F. J. (2019). Finding myself abroad: Adolescent identity processes and sociocultural adaptation during intercultural transitions. *Journal of Cross-Cultural Psychology, 50*, 594-614. https://doi.org/10.1177/0022022119831785

Headey, B., & Wearing, A. (1989). Personality, life events, and subjective well-being: Toward a dynamic equilibrium model. *Journal of Personality and Social Psychology*, *57*, 731–739. https://doi.org/10.1037/0022-3514.57.4.731

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*, 1–55. https://doi.org/10.1080/10705519909540118

Jackson, J. J., Thoemmes, F., Jonkmann, K., Lüdtke, O., & Trautwein, U. (2012). Military training and personality trait development: Does the military make the man, or does the man make the military? *Psychological Science*, *23*, 270–277. https://doi.org/10.1177/0956797611423545

Jeronimus, B. F., Riese, H., Sanderman, R., & Ormel, J. (2014). Mutual reinforcement between neuroticism and life experiences: A five-wave, 16-year study to test reciprocal causation. *Journal of Personality and Social Psychology*, *107*, 751–764. https://doi.org/10.1037/a0037009

Kandler, C., Bleidorn, W., Riemann, R., Angleitner, A., & Spinath, F. M. (2012). Life events as environmental states and genetic traits and the role of personality: A longitudinal twin study. *Behavior Genetics*, *42*, 57–72. https://doi.org/10.1007/s10519-011-9491-0

Kandler, C., & Ostendorf, F. (2016). Additive and synergetic contributions of neuroticism and life events to depression and anxiety in women. *European Journal of Personality, 30*, 390–405. https://doi.org/10.1002/per.2065

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*, 231–243. https://doi.org/10.1002/per.1952

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association, 90*, 773–795. Retrieved from https://www.stat.washington.edu/raftery/Research/PDF/kass1995.pdf

Kelley, K., & Rausch, J. R. (2006). Sample size planning for the standardized mean difference: Accuracy in parameter estimation via narrow confidence intervals. *Psychological Methods, 11*, 363–385. https://doi.org/10.1037/1082-989X.11.4.363

Kelloway, E. K. (2015). *Using Mplus for Structural Equation Modeling: A Researcher's Guide* (2nd edition). Sage Publications.

Kenny, D. A., & Judd, C. M. (2019). The unappreciated heterogeneity of effect sizes: Implications for power, precision, planning of research, and replication. *Psychological Methods*, *24*, 578–589. https://doi.org/10.1037/met0000209

King, R., & Ruiz-Gelices, E. (2003). International student migration and the European year abroad? Effects on European identity and subsequent migration behaviour. *International Journal of Population Geography*, *9*, 229–252. https://doi.org/10.1002/ijpg.280

Krings, F., Bangerter, A., Gomez, V., & Grob, A. (2008). Cohort differences in personal goals and life satisfaction in young adulthood: Evidence for historical shifts in developmental tasks. *Journal of Adult Development*, *15*, 93–105. https://doi.org/10.1007/s10804-008-9039-6

Lang, F. R., Lüdtke, O., & Asendorpf, J. B. (2001). Testgüte und psychometrische Äquivalenz der deutschen Version des Big Five Inventory (BFI) bei jungen, mittelalten und alten Erwachsenen [Validity and psychometric equivalence of the German version of the Big Five Inventory in young, middle-aged and old adults]. *Diagnostica*, *47*, 111–121. https://doi.org/10.1026//0012-1924.47.3.111

Le, K., Donnellan, M. B., & Conger, R. (2014). Personality development at work: Workplace conditions, personality changes, and the corresponsive principle. *Journal of Personality*, *82*, 44–56. https://doi.org/10.1111/jopy.12032

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. https://doi.org/10.1207/S15328007SEM0902_1

Luan, Z., Hutteman, R., Denissen, J. J.A., Asendorpf, J. B., & van Aken, M. A.G. (2017). Do you see my growth? Two longitudinal studies on personality development from childhood to young adulthood from multiple perspectives. *Journal of Research in Personality*, *67*, 44–60. https://doi.org/10.1016/j.jrp.2016.03.004

Lubke, G. H., Campbell, I., McArtor, D., Miller, P., Luningham, J., & van den Berg, S. M. (2017). Assessing model selection uncertainty using a bootstrap approach: An update. *Structural Equation Modeling: A Multidisciplinary Journal*, *24*, 230–245, https://doi.org/10.1080/10705511.2016.1252265

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*, 620–637. https://doi.org/10.1037/a0023743

Luhmann, M., Hofmann, W., Eid, M., & Lucas, R. E. (2012). Subjective well-being and adaptation to life events: A meta-analysis. *Journal of Personality and Social Psychology*, *102*, 592–615. https://doi.org/10.1037/a0025948

Luhmann, M., Orth, U., Specht, J., Kandler, C., & Lucas, R. E. (2014). Studying changes in life circumstances and personality: It's about time. *European Journal of Personality*, *28*, 256–266. https://doi.org/10.1002/per.1951

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

Marcusson-Clavertz, D., & Kjell, O. N. E. (2018). Psychometric properties of the spontaneous and deliberate mind wandering scales. Retrieved from https://econtent.hogrefe.com/doi/10.1027/1015-5759/a000470

McArdle, J. J., & Hamagami, F. (2002). Latent difference score structural models for linear dynamic analyses with incomplete longitudinal data. In L. M. Collins & A. Sayer (Eds.), *Decade of behavior. New methods for the analysis of change: Based on a conference held in 1998 at The Pennsylvania State University, a follow-up to the Los Angeles conference Best Methods for the Analysis of Change* (2nd ed., pp. 139–175). Washington, DC: American Psychological Association. https://doi.org/10.1037/10409-005

McCrae, R. R., Costa, P. T., de Lima, M. P., Simões, A., Ostendorf, F., Angleitner, A., … Piedmont, R. L. (1999). Age differences in personality across the adult life span: Parallels in five cultures. *Developmental Psychology, 35*, 466–477. https://doi.org/10.1037/0012-1649.35.2.466

McCrae, R. R., Costa, P. T., Ostendorf, F., Angleitner, A., Hřebíčková, M., Avia, M. D.,…Smith, P. B. (2000). Nature over nurture: Temperament, personality, and life span development. *Journal of Personality and Social Psychology, 78*, 173–186. https://doi.org/10.1037/0022-3514.78.1.173

Morris, S. B., & DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods, 7,* 105–125. https://doi.org/10.1037//1082-989X.7.1.105

Mõttus, R. (2017). Five-Factor Theory and personality development. In J. Specht (Ed.), *Personality development across the lifespan* (pp. 101–115). Elsevier Academic Press. https://doi.org/10.1016/B978-0-12-804674-6.00007-7

Muthén, L. K., & Muthén, B. O. (1998–2017). *Mplus User's Guide: Eigth Edition.* Los Angeles, CA: Muthén & Muthén. Retrieved from http://www.statmodel.com/download/usersguide/MplusUserGuideVer_8.pdf

Netz, N., & Jaksztat, S. (2014). Mobilised by mobility? Determinants of international mobility plans among doctoral candidates in Germany. In M. Tight & N. Maadad (Eds.), *International perspectives on higher education research: v. 11. Academic mobility* (Vol. 11, pp. 35–59). Bingley, U.K: Emerald. https://doi.org/10.1108/S1479-362820140000011009

Neyer, F. J., & Lehnart, J. (2007). Relationships matter in personality development: Evidence from an 8-year longitudinal study across young adulthood. *Journal of Personality, 75*, 535–568. https://doi.org/10.1111/j.1467-6494.2007.00448.x

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

Niehoff, E., Petersdotter, L., & Freund, P. A. (2017). International sojourn experience and personality development: Selection and socialization effects of studying abroad and the Big Five. *Personality and Individual Differences*, *112*, 55–61. https://doi.org/10.1016/j.paid.2017.02.043

Reuter, T., Ziegelmann, J. P., Wiedemann, A. U., Geiser, C., Lippke, S., Schüz, B., & Schwarzer, R. (2010). *Journal of Health Psychology, 15*, 935–947. https://doi.org/10.1177/1359105309360071

Rivza, B., & Teichler, U. (2007). The changing role of student mobility. *Higher Education Policy*, *20*, 457–475. https://doi.org/10.1057/palgrave.hep.8300163

Roberts, B. W. (2018). A revised sociogenomic model of personality traits. *Journal of Personality, 86*, 23–35. https://doi.org/10.1111/jopy.12323

Roberts, B. W., Caspi, A., & Moffitt, T. E. (2003). Work experiences and personality development in young adulthood. *Journal of Personality and Social Psychology*, *84*, 582–593. https://doi.org/10.1037/0022-3514.84.3.582

Roberts, B. W., Hill, P. L., & Davis, J. P. (2017). How to change conscientiousness: The sociogenomic trait intervention model. *Personality Disorders, 8*, 199–205. https://doi.org/10.1037/per0000242

Roberts, B. W., & Jackson, J. J. (2008). Sociogenomic personality psychology. *Journal of Personality, 76*, 1523–1544. https://doi.org/10.1111/j.1467-6494.2008.00530.x

Roberts, B. W., Luo, J., Briley, D. A., Chow, P. I., Su, R., & Hill, P. L. (2017). A systematic review of personality trait change through intervention. *Psychological Bulletin, 143*, 117–141. https://doi.org/10.1037/bul0000088

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

Roberts, B. W., Wood, D., & Smith, J. L. (2005). Evaluating Five Factor theory and social investment perspectives on personality trait development. *Journal of Research in Personality, 39*, 166–184. https://doi.org/10.1016/j.jrp.2004.08.002

Rodríguez González, C., Bustillo Mesanza, R., & Mariel, P. (2011). The determinants of international student mobility flows: An empirical study on the Erasmus programme. *Higher Education, 62*, 413–430. https://doi.org/10.1007/s10734-010-9396-5

Salisbury, M. H., Umbach, P. D., Paulsen, M. B., & Pascarella, E. T. (2009). Going global: Understanding the choice process of the intent to study abroad. *Research in Higher Education, 50*, 119–143. https://doi.org/10.1007/s11162-008-9111-x

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods, 7*, 147–177. https://doi.org/10.1037/1082-989X.7.2.147

Specht, J., Bleidorn, W., Denissen, J. J. A., Hennecke, M., Hutteman, R., Kandler, C., . . . Zimmermann, J. (2014). What drives adult personality development? A comparison of theoretical perspectives and empirical evidence. *European Journal of Personality, 28*, 216–230. https://doi.org/10.1002/per.1966

Specht, J., Egloff, B., & Schmukle, S. C. (2011). Stability and change of personality across the life course: The impact of age and major life events on mean-level and rank-order stability of the Big Five. *Journal of Personality and Social Psychology, 101*, 862–882. https://doi.org/10.1037/a0024950

Steyer, R., Eid, M., & Schwenkmezger, P. (1997). Modeling true intraindividual change: True change as a latent variable. *Methods of Psychological Research Online, 2*(1), 21–33. Retrieved from https://www.dgps.de/fachgruppen/methoden/mpr-online/issue2/art2/steyer.pdf

Steyer, R., Mayer, A., Geiser, C., & Cole, D. A. (2015). A theory of states and traits—Revised. *Annual Review of Clinical Psychology, 11*, 71–98. https://doi.org/10.1146/annurev-clinpsy-032813-153719

Steyer, R., Partchev, I., & Shanahan, M. J. (2000). Modeling true intradindividual change in structural equation models: The case of poverty and children's psychosocial adjustment. In T. D. Little, K. U. Schnabel, & J. Baumert (Eds.), *Modeling longitudinal and multilevel data: Practical issues, applied approaches, and specific examples* (109–126). Mahwah, N.J.: Lawrence Erlbaum Associates.

Suanet, I., & van de Vijver, F. J. R. (2009). Perceived cultural distance and acculturation among exchange students in Russia. *Journal of Community & Applied Social Psychology, 19*, 182–197. https://doi.org/10.1002/casp.989

Tamura, T., & Furnham, A. (1993). Comparison of adaptation to the home culture of Japanese children and adolescents returned from overseas sojourn. *International Journal of Social Psychiatry*, *39*, 10–21. https://doi.org/10.1177/002076409303900102

van de Schoot, R., Lugtig, P., & Hox, J. (2012). A checklist for testing measurement invariance. *European Journal of Developmental Psychology*, *9*, 486–492. https://doi.org/10.1080/17405629.2012.686740

Wagner, J., Orth, U., Bleidorn, W., Hopwood, C. J., & Kandler, C. (2020). Towards an integrative model of sources of personality stability and change. *Current Directions in Psychological Science*.

Ward, C., Okura, Y., Kennedy, A., & Kojima, T. (1998). The U-Curve on trial: A longitudinal study of psychological and sociocultural adjustment during cross-cultural transition. *International Journal of Intercultural Relations, 22*, 277–291. https://doi.org/10.1016/S0147-1767(98)00008-X

Wrzus, C., & Roberts, B. W. (2017). Processes of personality development in adulthood: The TESSERA framework. *Personality and Social Psychology Review: An Official Journal of the Society for Personality and Social Psychology, Inc*, *21*, 253–277. https://doi.org/10.1177/1088868316652279

Zimmermann, J., Greischel, H., & Jonkmann, K. (2020). The development of multicultural effectiveness in international student mobility. *Higher Education*. https://doi.org/10.1007/s10734-020-00509-2

Zimmermann, J., & Neyer, F. J. (2013). Do we become a different person when hitting the road? Personality development of sojourners. *Journal of Personality and Social Psychology*, *105*, 515–530. https://doi.org/10.1037/a0033019

Zimmermann, J., Schubert, K., Bruder, M., & Hagemeyer, B. (2017). Why going the extra mile? A longitudinal study on sojourn goals and their impact on sojourners' adaptation. *International Journal of Psychology. 52*, 425–435. https://10.1002/ijop.12240