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

JULIA RICHTER[1]*, JULIA ZIMMERMANN[2], FRANZ J. NEYER[3] and CHRISTIAN KANDLER[4]

[1]*Department of Psychology, Bielefeld University, Bielefeld, Germany*
[2]*Department of Psychology, FernUniversität in Hagen, Hagen, Germany*
[3]*Department of Psychology, Friedrich-Schiller University Jena, Jena, Germany*
[4]*Department of Psychology, University of Bremen, Bremen, Germany*

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 = 1095) 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 neighbour-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. Because of 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. © 2020 The Authors. European Journal of Personality published by John Wiley & Sons Ltd on behalf of European Association of Personality Psychology

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

INTRODUCTION

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 behavioural 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 with non-sojourners' traits several years after their participation in international student mobility.

*Correspondence to: Julia Richter, Department of Psychology, Bielefeld University, PO box 100131, 33501 Bielefeld, Germany.
E-mail: julia.richter1@uni-bielefeld.de

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, for example 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; Lüdtke et al., 2011; Luhmann, Hofmann, Eid, & Lucas, 2012; Neyer & Lehnart, 2007; Specht et al., 2014), even within a short period of time (Roberts 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).
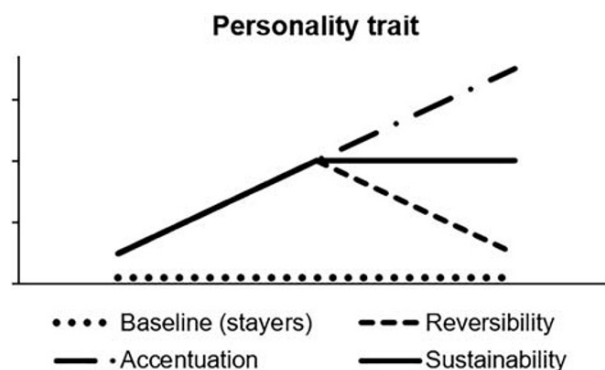

Figure 1. Examples of long-term trait change patterns as a result of sojourn experiences in comparison to stayers' baseline changes.

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 with 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, that is 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 behaviour. For example, Tina found herself with a lot of new behaviours and ideas because of 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

[1]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.

steeply (or increases less steeply) than her friends', leading to decreased mean-level differences between them after her return as compared with the previous measurement occasion(s) (see Figure 1). One reason for a reversed effect might be that Tina picks up old habits and behaviours, and adapts again to her home's norms (Denissen, van Aken, Penke, & Wood, 2013).[1]

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 because of 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 behaviours. 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, e.g. 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 because of 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 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 with 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 analysing 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/).[2] The current investigation went beyond the previous study by adding a follow-up measurement occasion (T4) about five years after the first measurement took place. All measurements were carried out using online questionnaires. During the first study period, a sample of $N = 1134$ German university students were repeatedly tracked over the course of the academic year 2009/2010 (T1, T2, and T3). Data were collected two weeks before (T1), six months after (T2; time range: four to eight 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 2009/2010.

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 ($N = 1134$), we excluded $n = 39$ participants who had reported their main place of residence at T4 outside of Germany. From the resulting total sample of $n = 1095$ students, $n = 498$ were *sojourners* who went abroad shortly after the start of the study at T1. The vast

---

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

majority of the sojourners, 84.9%, moved to a European country ($n = 423$), 6.0% ($n = 30$) to Asia, 5.4% ($n = 27$) to the USA 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 2009/2010 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 2009/2010 ($n = 597$). Mean age at survey start was 22.59 years ($SD = 2.60$, age range: 18–41 years), and 78.0% of the participants were female ($n = 854$).

Drop-out

At T3, $n = 1059$ (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 with 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 test statistics. Little's Missing-Completely-at-Random 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 programme 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, and parental leave.[3]

Measures

Personality trait measures

We applied the German version of the Big Five Inventory (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 2009/2010 (stayers) and participants with sojourn experiences during the academic year 2009/2010 (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, that is 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, that is 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 2009/2010, 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 and 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 neighbour-change models. For both approaches, were carried out univariate as well as multivariate

---

[3]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

Top panel — columns T1 and T2

	T1 O	T1 C	T1 E	T1 A	T1 N	T2 O	T2 C	T2 E	T2 A	T2 N
T1										
O	*.83*					.86/.85	.10/.03	.26/.23	.07/.03	−.08/−.01
C	.11/.03/.19	*.82*				.08/.05	.85/.87	.21/.19	.27/.26	−.14/−.11
E	.29/.27/.31	.19/.11/.26	*.8*			.25/.20	.21/.17	.87/.88	.18/.16	−.30/−.26
A	.09/.02/.16	.32/.25/.36	.16/.02/.26	*.72*		.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

Bottom panel — columns T3 and T4

	T3 O	T3 C	T3 E	T3 A	T3 N	T4 O	T4 C	T4 E	T4 A	T4 N
T1										
O	.84/.82	.10/.04	.26/.24	.06/.03	−.09/−.01	.71/.72	−.02/−.02	.18/.18	.04/.06	−.01/.03
C	.07/.04	.85/.86	.21/.20	.26/.28	−.11/−.10	.02/.05	.75/.77	.16/.14	.26/.23	−.06/−.07
E	.23/.20	.22/.19	.86/.87	.17/.16	−.28/−.24	.13/.11	.21/.18	.78/.74	.13/.15	−.16/−.18
A	.07/.01	.31/.29	.20/.18	.78/.78	−.25/−.23	.02/.03	.24/.26	.14/.13	.66/.67	−.29/−.28
N	−.09/−.02	−.19/−.18	−.32/−.31	−.28/−.30	.80/.81	.00/.02	−.10/−.10	−.22/−.19	−.21/−.23	.69/.68
T2										
O	.88/.87	.12/.05	.27/.22	.10/.06	−.10/−.02	.74/.73	−.00/.00	.16/.14	.07/.05	.00/.05
C	.09/.03	.87/.88	.25/.21	.29/.30	−.16/−.14	.01/.04	.75/.78	.18/.14	.28/.27	−.09/−.11
E	.27/.23	.26/.23	.90/.90	.23/.19	−.33/−.28	.14/.14	.23/.19	.77/.76	.17/.17	−.18/−.18
A	.10/.03	.32/.30	.24/.20	.83/.83	−.34/−.31	.03/.01	.21/.23	.12/.09	.68/.69	−.31/−.28
N	−.09/−.02	−.21/−.19	−.36/−.32	−.32/−.29	.85/.85	.01/.00	−.12/−.10	−.25/−.22	−.28/−.27	.73/.75
T3										
O	*.84*	.12/.05	.29/.26	.10/.04	−.12/−.04	.73/.70	.01/−.01	.19/.16	.07/.02	.02/.09
C	.18/.06/.27	*.83*	.29/.20/.34	.34/.33	−.21/−.20	.02/.05	.75/.79	.19/.17	.30/.29	−.12/−.15
E	.29/.26/.33	.29/.20/.34	*.89*	.27/.18/.33	−.36/−.33	.15/.17	.24/.23	.79/.76	.16/.17	−.20/−.20
A	.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*

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

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

Above the diagonal: Total sample[†]/stayers[‡]. *Below the diagonal:* All sojourners[§]/short-term sojourners[¶]/long-term sojourners[††].

[†]$N_{T1–T3} = 1,095$, $n_{T4} = 441$.
[‡]$N_{T1–T3} = 597$, $n_{T4} = 239$.
[§]$N_{T1–T3} = 498$, $N_{T4} = 202$.
[¶]$N_{T1–T3} = 218$, $n_{T4} = 88$.
[††]$N_{T1–T3} = 280$, $n_{T4} = 114$.

analyses. Although we examined results based on listwise deletion and full information maximum likelihood (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 neighbour-change modelling approach in more detail. We first reflect on the latent variable and latent change modelling, 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 modelling

We modelled 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 Table S1. For the latent analyses, all latent factor indicators (i.e. parcels) were standardized.

Latent change modelling

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 analysed personality trait changes across T1–T4 only. Thereafter, we analysed all four time points using latent neighbour-change models to capture (discontinuous) change during each time interval (i.e. from T1 to T2, from T2 to T3, and from T3 to T4; Steyer, Partchev, & Shanahan, 2000). In latent neighbour-change models, latent trait change reflects the difference between latent trait levels at two neighbouring 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

[4]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/.

Table 2. Group-based descriptive statistics

Group	M (SD)				Effect size for repeated measures[†]			
	T1	T2	T3	T4	d_{12}	d_{23}	d_{34}	d_{14}
Full sample[‡]								
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[§]								
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[¶]								
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[††]								
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[‡‡]								
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: A, agreeableness; C, conscientiousness; E, extraversion; N, neuroticism; O, openness.

[†]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.

[‡]$N_{T1-T3} = 1{,}095$; $n_{T4} = 441$.

[§]$N_{T1-T3} = 597$, $n_{T4} = 239$.

[¶]$N_{T1-T3} = 498$, $n_{T4} = 202$.

[††]$N_{T1-T3} = 218$, $n_{T4} = 88$.

[‡‡]$N_{T1-T3} = 280$, $n_{T4} = 114$.

loadings, intercepts, and residual variances of the same parcels per trait across time (Marcusson-Clavertz & Kjell, 2018).[4]

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 procedures). Loadings of the respective parcels on the indicator-specific factor were modelled 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 2009/2010 on trait changes between T1 and T2, T2 and T3, and between T3 and T4.

Multivariate trait change modelling

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
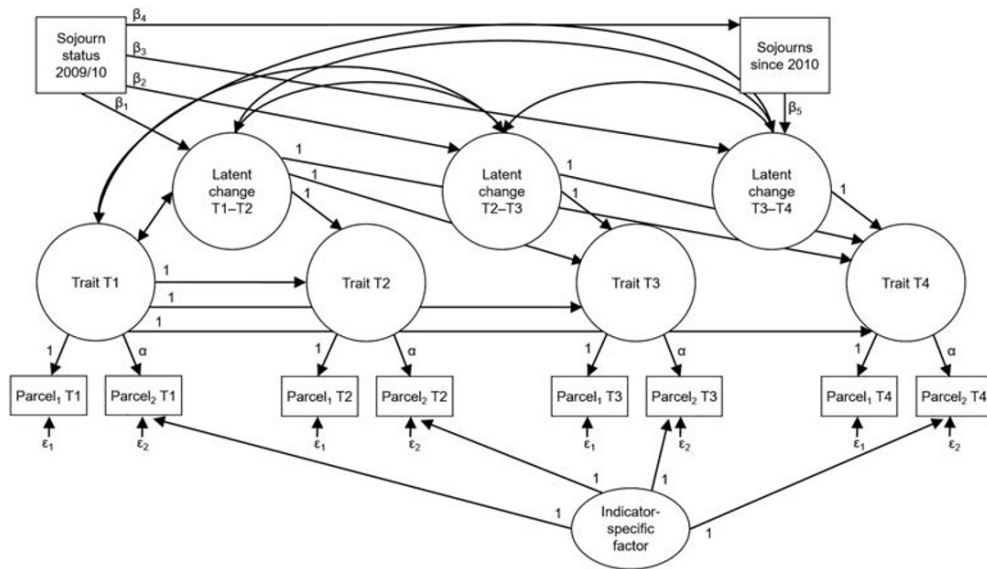
Figure 2. Conceptual model of sojourn effects on trait changes across T1–T2, T2–T3, and T3–T4. For reasons of parsimony, we only printed an exemplary univariate latent neighbour-change model, although we also imputed multivariate latent neighbour-change models. We controlled for age, sex, and sojourns 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.

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).[5]

Sojourn effects modelling
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 2009/2010 on trait changes between T3 and T4 after controlling for prior effects 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 2009/2010 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 2009/2010 on further sojourn experiences during the T3–T4 interval.

Analytical criteria
We evaluated the absolute goodness of model fit using the root mean square error of approximation (RMSEA), comparative fit index (CFI), and standardized root mean square residual (SRMR). RMSEA portrays the average proportion

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

of misfit per degree of freedom; CFI is an indicator for the relative noncentrality 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 the 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 FIML method, which is a model-based approach to handle missing data (Schafer & Graham, 2002). In comparison with 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 analysed 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 with stayers, sojourners tended to show higher levels in openness, conscientiousness, extraversion, agreeableness, and emotional stability (vs.
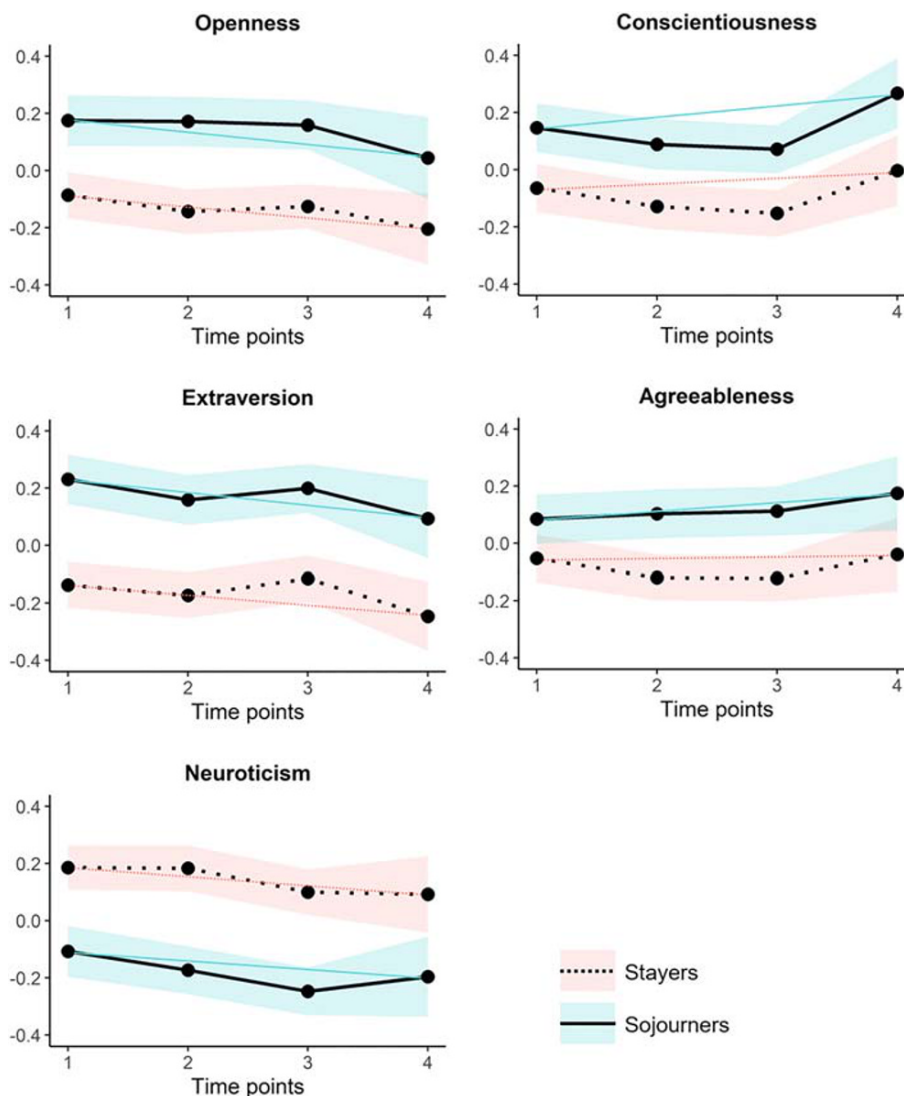
Figure 3. Descriptive mean-level differences of stayers and sojourners based on raw means of manifest personality trait scores. We plotted the raw means standardized around the grand mean of each trait to make the figures more comparable. Estimates include group-based 95% confidence intervals to illustrate measurement inaccuracy. We included a linear trend line between T1 and T4 in each groups' respective colour 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 because of 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. [Colour figure can be viewed at wileyonlinelibrary.com]

neuroticism). Moreover, these differences tended to be constant across all four measurement occasions. See also 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 because of unconsidered intertrait 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 neighbour-change models controlling for potential biases. Please refer to Figure S2 for time-specific boxplots of the three groups based on listwise deletion.

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 Supporting information at the open science framework (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 latent model and the multivariate latent model showed good fits to the data (see 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 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 with 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 neighbour-change models across four time points fit well to the data (see 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.
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 among sojourners compared with 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,

Table 3. Sojourn effects on trait changes as derived from the latent-change analyses across T1–T4 only

Type of analysis	Sojourn at T1						Sojourn between T3 and T4					
	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 = 1095. b = unstandardized effect estimate. Sojourn at T1 = sojourn in the academic term 2009/2010; Sojourn between T3 and T4 = sojourns since 1 October 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. CI, confidence interval.
For model fit indices, see Table S2.
Significant model parameters (p < .05) are shown in boldface.

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 and T4→															
ΔT3–T4	.08	.079	−.01, .17	.07	.101	−.01, .16	.05	.271	−.04, .13	.05	.291	−.04, .13	−.03	.401	−.11, .05

Extraversion			Agreeableness						Neuroticism					
Multivariate			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

Note: $N = 1095$. b = unstandardized effect estimate. Sojourn at T1 = sojourn in the academic term 2009/2010; Sojourn between T3 and T4 = sojourns since 1 October 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. CI, confidence interval.
For model fit indices, see supplementary Table S2.
Significant model parameters ($p < .05$) are shown in boldface.

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 Table S4). As point estimates were quite comparable across FIML and listwise deletion, this divergence in statistical significance was most likely because of the lower statistical power of the smaller listwise-deletion sample. Accordingly, the CIs revealed higher accuracy in the FIML-based analyses.

Figure 4 is based on the latent modelling 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 that accelerated neuroticism decrease in sojourners, with no substantial change in stayers. Please note that Figures 3 and 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-term 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 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 modelling 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 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' neuroticism after their return back home compared with the other groups (see 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-term 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).
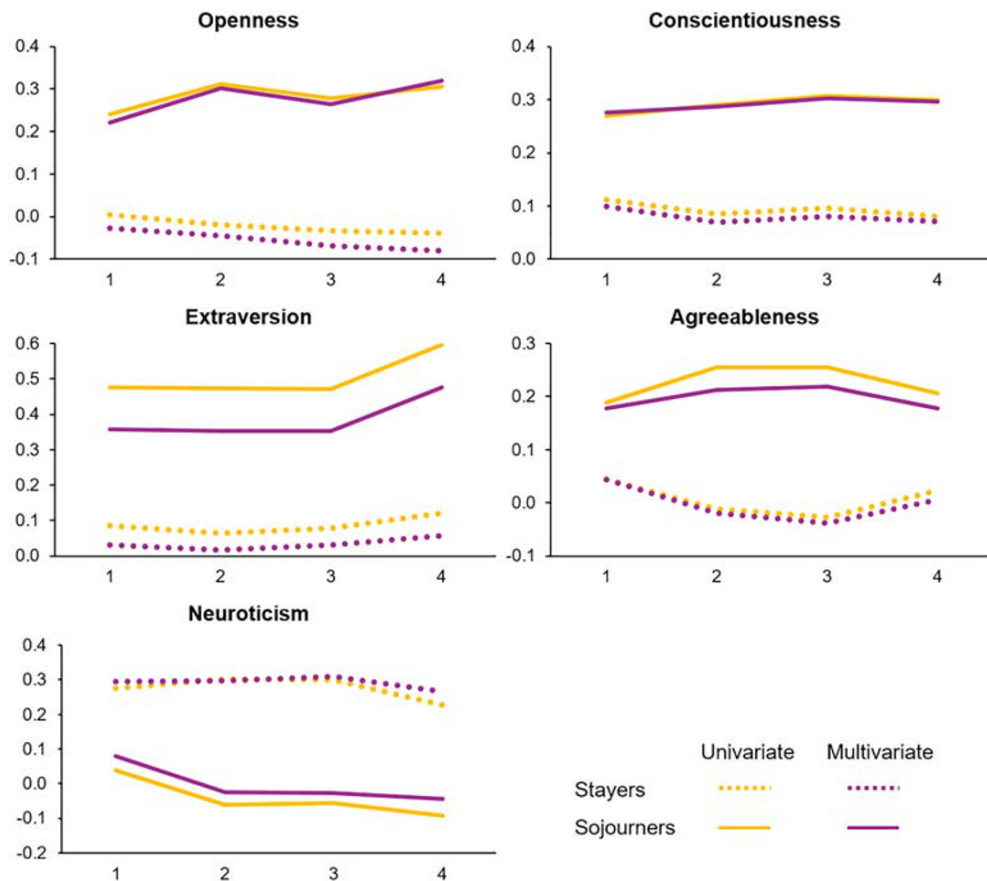
Figure 4. Big Five latent change trends of stayers and sojourners. We printed the respective estimates derived from the univariate and multivariate latent neighbour-change models after standardizing the latent factor indicators, centred on age and sex. T1 values are the unweighted means of the estimates that derived from additional regressions of the Big Five traits on sojourn status in our models, also centred 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. [Colour figure can be viewed at wileyonlinelibrary.com]

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 (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 Table S5), 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 because of 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 modelling 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 with T1–T2. Accordingly, there were much wider 95% CIs for T3–T4 estimates compared with 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

Table 5. Results of investigating long-term sojourn effects more deeply

Statistics	Openness					Conscientiousness					Extraversion				
	Wald tests			Bayesian criteria		Wald tests			Bayesian criteria		Wald tests			Bayesian criteria	
	$\Delta\chi^2$	df	p	BIC	SABIC	$\Delta\chi^2$	df	p	BIC	SABIC	$\Delta\chi^2$	df	p	BIC	SABIC
Univariate															
Rev.	3.04	1	.081	14 351.06	14 255.78	1.00	1	.318	15 214.11	15 118.82	2.53	1	.112	13 350.63	13 255.34
Accent.	1.14	1	.286	14 349.17	14 253.88	0.41	1	.522	15 213.53	151 18.24	0.86	1	.354	13 348.96	13 253.67
Sust.	0.12	1	.735	14 348.15†	14 252.86†	0.03	1	.858	15 213.15	15 117.86	1.59	1	.208	13 349.68	13 254.40
Multivariate															
Rev.	**5.92**	1	.015	71 216.16	70 193.41	0.50	1	.482	71 210.77	70 188.03	2.84	1	.092	71 213.11	70 190.36
Accent.	0.38	1	.539	71 210.65†	70 187.91†	0.34	1	.559	71 210.62	70 187.87	1.84	1	.175	71 212.11	70 189.36
Sust.	0.83	1	.364	71 211.10†	70 188.36†	0.00	1	.952	71 210.28	70 187.53	2.31	1	.129	71 212.58	70 189.84

	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	16 863.86†	16 768.58†	1.88	1	.170	15 636.90†	15 541.61†
	10.02	1	.002	16 873.60	16 778.31	**6.76**	1	.009	15 641.73	15 546.45
	1.88	1	.170	16 865.57†	16 770.28†	0.38	1	.539	15 635.40†	15 540.11†
	0.15	1	.695	71 210.43†	70 187.68†	1.60	1	.206	71 211.87	70 189.13
	6.45	1	.011	71 216.68	70 193.94	3.08	1	.079	71 213.35	70 190.60
	1.15	1	.283	71 211.43†	70188.68†	0.06	1	.806	71 210.34†	70 187.59†

Note: N = 1095. Rev., reversed effects: setting the effect between T3 and T4 at the opposite of the respective effect between T1 and T2; Accent., accentuated effects: setting the effect between T3 and T4 at the same as the respective effect between T1 and T2; Sust., sustained effects: setting the effect for T3–T4 to 0. BIC, Bayesian information criterion; SABIC, sample-size adjusted Bayesian information criterion.
For model fit indices, see 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.

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. $b_{\text{univariate}}$ = −.09 and $b_{\text{multivariate}}$ = −.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*).[6] For example, we tested whether equating the sojourn effect on openness change T3–T4 to the effect on openness change T1–T2 (i.e. $b_{\text{univariate}}$ = .09 and $b_{\text{multivariate}}$ = .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,[7] we 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

[6]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 six-months interval from T1 to T2 and the four years between T3 and T4 was deemed the most conservative prognosis.

[7]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.

sojourn effects tended to be unlikely for openness, but not for the other traits.

Accentuated effects under investigationWald 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 investigationWe 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 effectsThe 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 the SABIC, this criterion penalizes sample size less (Kelloway, 2015), but is comparable with 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 ($\Delta BIC_{reversed-accentuated, \ multivariate} = 5.51$), but not in the univariate model ($\Delta BIC_{reversed-accentuated, \ univariate} = 1.89$). Comparing the sustained-effect model to the accentuated-effect model yielded no substantial difference between these models for openness ($\Delta BIC_{accentuated-sustained, \ univariate/multivariate} = 1.02/-0.45$), while comparing it to the reversed-effect model indicated a better fit for the sustained-effect model ($\Delta BIC_{reversed-sustained, \ univariate/multivariate} = 2.91/5.06$). 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-effect or accentuated-effect model ($\Delta BIC_{reversed-accentuated, \ univariate/multivariate} = 0.58/0.15$) and the sustained-effect model ($\Delta BIC_{reversed-sustained, \ univariate/multivariate} = 0.96/0.49$, and $\Delta BIC_{accentuated-sustained, \ univariate/multivariate} = 0.38/0.34$). For extraversion ($\Delta BIC_{reversed-accentuated, \ univariate/multivariate} = 1.67/1.00$, $\Delta BIC_{reversed-sustained, \ univariate/multivariate} = 0.95/0.53$, and $\Delta BIC_{accentuated-sustained, \ univariate/multivariate} = -0.72/-0.47$), 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 ($\Delta BIC_{reversed-accentuated, \ univariate/multivariate} = -9.74/-6.25$) and a sustained-effect model ($\Delta BIC_{accentuated-sustained, \ univariate/multivariate} = 8.03/5.25$) both fit better than an accentuated-effect model. There was no substantial difference between the reversed-effect and the sustained-effect model ($\Delta BIC_{reversed-sustained, \ univariate/multivariate} = -1.71/-1.00$). For neuroticism, the sustained-effect model ($\Delta BIC_{accentuated-sustained, \ univariate/multivariate} = 6.33/3.01$) fit better than the accentuated-effect model. While the results were indifferent for comparing the reversed-effect and the accentuated-effect model ($\Delta BIC_{reversed-accentuated, \ univariate/multivariate} = -4.83/-1.48$), reversed-effect and sustained-effect models did not differ markedly ($\Delta BIC_{reversed-sustained, \ univariate/multivariate} = 1.50/1.53$). 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 neighbour-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 ($b_{univariate} = .04$, $b_{multivariate} = .02$) than for T1–T2 ($b_{univariate} = -.13$, $b_{multivariate} = -.10$). 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 ($b_{univariate} = .12$, $b_{multivariate} = .10$). The opposite trend ($b_{univariate} = -.10$,

$b_{multivariate} = -.09$) 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 CIs in Table 4 (95% CI$_{univariate}$ [−.23, .04], 95% CI$_{multivariate}$ [−.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 favour 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 2009/2010 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 female participants tended to decrease more strongly or increased less strongly in openness than male participants, 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 female paticipants' openness decreased more or increased less than male participants' 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 female participants' 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 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 with 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 with 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

Table 6. Age and sex effects on trait changes as derived from the latent-change analyses across T1–T4 only

Effect	Openness			Conscientiousness			Extraversion			Agreeableness			Neuroticism		
	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 = 1095. 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 centred before the analysis. CI, confidence interval. For model fit indices, see 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

Extraversion			Agreeableness						Neuroticism					
Multivariate			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

Note: N = 1095. 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 centred before the analysis.
For model fit indices, see Table S2.
Significant model parameters (*p* < .05) are shown in boldface.

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 & Jackson, 2008; Roberts, 2018).

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 four to six 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 from both 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 CIs 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 towards a reversed sojourn effect, the point estimate based on the latent neighbour-change models did not clearly support this suggestion as the coefficient was close to zero (see Table 4, and compare Figures 3 and 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 with 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 modelling
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 intertrait 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 with the significant effects found for the T1–T2 interval, but were not statistically significant because of 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 was the lower power and accuracy to detect effects for the T1–T4 and the T3–T4 intervals because of 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 nonexperimental 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 because of 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 and 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. Other further alternatives 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 behaviour (changes) (e.g. with ambulatory assessment methods) might be a promising way to assess and integrate the short-term and long-term processes of personality trait change (see Geukes, van Zalk, & Back, 2018; Wrzus & Roberts, 2017). With regard to the microprocesses 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, female participants 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, that is by engaging in new behaviours 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 behaviour might have led sojourners to change their habits and daily behaviours throughout their sojourn. Now, several years later, some of these shifted values might still be in place and regulate their day-to-day behaviour, 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-ratings 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 predeparture 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, that is 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 and 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.

The present research partly used data from the longitudinal study 'PEDES—Personality Development of Sojourners', which has been described by Zimmermann and Neyer (2013) and Zimmermann & Neyer et al. (2014). The current study differs from the previous publications in the following aspects. First, it focuses on another research question (i.e. the sustainability of sojourn effects on personality traits across five years). Second, it incorporates data from a new measurement occasion (i.e. a follow-up measurement that was taken five years after the end of the previous study period). Hence, there is no substantial overlap with previous publications based on the PEDES data.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table S1. Parcel-Based Reliability Coefficients of Personality Traits
Table S2. *Comparing Strong to Strict Measurement Invariance*
Table S3. *Sojourn Effects on Trait Changes as Derived from the Latent-Change Analyses Across T1–T4 Only Based on Listwise Deletion*
Table S4. *Sojourn Effects on Trait Changes as Derived from the Latent-Change Analyses Across Four Time Points Based on Listwise Deletion*
Table S5. *Short and Long-term Sojourn Effects on Trait Changes as Derived from the Latent-Change Analyses Across Four Time Points.*
Table S6. *Model Fit Indices After Constraining Long-Term Sojourn Effects*
Figure S1. Descriptive mean-level differences of stayers, short-term sojourners, and long-term 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 standard errors to illustrate measurement inaccuracy. We included a linear trend line between T1 and T4 in each groups' respective color theme to make 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
Figure S2. Box plots of raw means arranged by group and time point. We calculated the estimates based on listwise deletion of cases across all time points. This led to results diverging from Figures 3 and S1.
Figure S3. Latent extraversion and neuroticism change trends of stayers, short-term sojourners, and long-term 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 centred 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.

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.*, *34*, 285–300. 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, *86*, 83–96. 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*, *34*, 301–321. 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, https://doi.org/10.1080/01621459.1995.10476572

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 ed.). 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 correspondive 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.

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.

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.

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). Changes in intentions, planning, and self-efficacy predict changes in behaviors. *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*, 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 intraindividual 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* (pp. 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) (accepted)). 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://doi.org/10.1002/ijop.12240