The transition to grandparenthood: No consistent evidence for change in the Big Five personality traits and life satisfaction

Michael D Krämer1,2,3, Manon A van Scheppingen4, William J. Chopik5 and David Richter1,2,3

Abstract

Intergenerational relations have received close attention in the context of population aging and increased childcare provision by grandparents. However, few studies have investigated the psychological consequences of becoming a grandparent. In a preregistered test of grandparenthood as a developmental task in middle and older adulthood, we used representative panel data from the Netherlands (N = 563) and the United States (N = 2210) to analyze first-time grandparents’ personality and life satisfaction development. We tested gender, employment, and grandchild care as moderators. To address confounding, we employed propensity score matching using two procedures: matching grandparents with parents and nonparents to achieve balance indifferent sets of carefully selected covariates. Multilevel models demonstrated mean-level stability of the Big Five personality traits and life satisfaction over the transition to grandparenthood, and no consistent moderation effects—contrary to the social investment principle. The few small effects of grandparenthood on personality development did not replicate across samples. We found no evidence of larger interindividual differences in change in grandparents compared to the controls or of lower rank-order stability. Our findings add to recent critical re-examinations of the social investment principle and are discussed in light of characteristics that might moderate grandparents’ personality development.

Keywords

Grandparenthood, Big Five, life satisfaction, personality development, propensity score matching

Received 28 February 2022; Revised 19 July 2022; accepted 28 February 2022

Introduction

Becoming a grandparent is an important life event for many people in midlife or old age (Infurna et al., 2020). In an era of population aging, the time that grandparents are alive and in good health is prolonged compared to previous generations (Bengtson, 2001; Leopold & Skopek, 2015; Margolis & Wright, 2017). In addition, grandparents fulfill an increased share of childcare responsibilities (Hayslip et al., 2019; Pilkauskas et al., 2020). In recent years, intergenerational relations have received heightened attention from psychological and sociological research (Bengtson, 2001; Coall & Hertwig, 2011; Fingerman et al., 2020). In research on personality development, the transition to grandparenthood has been proposed as an important developmental task arising in old age (Hutteman et al., 2014). However, empirical research on the psychological consequences of grandparenthood remains sparse. Using data from two nationally representative panel studies, we investigate whether the transition to grandparenthood affects the Big Five personality traits and life satisfaction. We test hypotheses derived from neo-socioanalytic theory (Roberts & Wood, 2006) in a prospective quasi-experimental case-control design (see Luhmann et al., 2014).
personality development takes place in adolescence and emerging adulthood (Bleidorn & Schwaba, 2017; Pusch et al., 2019; Schwaba & Bleidorn, 2018), personality traits also change in middle and older adulthood (e.g., Allemand et al., 2008; Damian et al., 2019; Kandler et al., 2015; Lucas & Donnellan, 2011; Mõttus et al., 2012; Mueller et al., 2016; Seifert et al., 2021; Wagner et al., 2016; for a review, see Specht, 2017).

Here, we examine the Big Five personality traits— agreeableness, conscientiousness, extraversion, neuroticism, and openness to experience—which constitute a broad categorization of universal patterns of thought, affect, and behavior (John et al., 2008; John & Srivastava, 1999). Changes over time in the Big Five occur both in mean trait levels (i.e., mean-level change; Roberts et al., 2006) and in the ordering of people relative to each other on trait dimensions (i.e., rank-order stability; Anusic & Schimmack, 2016; Roberts & DelVecchio, 2000). A lack of observed changes in mean trait levels does not necessarily mean that individual trait levels are stable over time, and perfect rank-order stability does not preclude mean-level changes. Mean-level changes in early to middle adulthood (circa 30–60 years old; Hutteman et al., 2014) are typically characterized by greater maturity, as evidenced by increased agreeableness and conscientiousness and decreased neuroticism (Damian et al., 2019; Roberts et al., 2006). In old age (circa 60 years and older; Hutteman et al., 2014), research is generally more sparse. But there is some evidence of a reversal of this positive personality development following retirement (sometimes termed la dolce vita effect; Asselmann & Specht, 2021; Marsh et al., 2013; cf. Schwaba & Bleidorn, 2019) and at the end of life when health problems arise (Wagner et al., 2016).

In terms of rank-order stability, most prior studies have shown support for an inverted-U shape trajectory (Ardelt, 2000; Lucas & Donnellan, 2011; Seifert et al., 2021; Specht et al., 2011; Wortman et al., 2012): Rank-order stability rises until it reaches a plateau in midlife, and decreases in old age. However, evidence is mixed on whether rank-order stability decreases again in old age (see Costa et al., 2019; Wagner et al., 2019). We are not aware of any study investigating trait rank-order stability over the transition to grandparenthood. Other life events are associated with rank-order stability of personality and well-being, although only certain events and traits (e.g., Denissen et al., 2019; Hentschel et al., 2017; Specht et al., 2011). Still, the previously held view that personality is stable or “set like plaster” (Specht, 2017, p. 64) after one reaches adulthood (or leaves emerging adulthood behind; Bleidorn & Schwaba, 2017) has been largely abandoned (Specht et al., 2014).

Theories explaining the mechanisms of personality development in middle and older adulthood emphasize genetic influences and life experiences as interdependent sources of stability and change (Bleidorn et al., 2021; Specht et al., 2014; Wagner et al., 2020). We conceptualize the transition to grandparenthood as adopting a new social role according to the social investment principle of neo-socioanalytic theory (Lodi-Smith & Roberts, 2007; Roberts & Wood, 2006). The social investment principle states that normative life events or transitions such as entering the workforce or becoming a parent lead to personality maturation through adopting new social roles (Roberts et al., 2005). These new roles encourage or compel people to act in a more agreeable, conscientious, and emotionally stable (i.e., less neurotic) way. People’s experiences in these roles as well as societal expectations towards them are hypothesized to drive long-term personality development (Lodi-Smith & Roberts, 2007; Wrzus & Roberts, 2017).

Empirical research on life events entailing new social roles has focused on young adulthood: A first romantic relationship (Wagner et al., 2015), the transition from high school to university, or a first job (Asselmann & Specht, 2021; Golle et al., 2019; Lüdtke et al., 2011) co-occur with mean-level changes that are (partly) consistent with the social investment principle (for a review, see Bleidorn et al., 2018). However, recent findings on the transition to parenthood fail to support the social investment principle (Asselmann & Specht, 2020b; van Scheppingen et al., 2016). An analysis of trajectories of the Big Five before and after different life events produced limited support for the social investment principle: Small increases in emotional stability occurred following the transition to employment but not in the other traits or following marriage or childbirth (Denissen et al., 2019).

Age-graded, normative role transitions may drive personality development across the entire lifespan but they are understudied in middle and older adulthood. Recent research indicates that retirement contributes to personality change following a period of relative stability in middle life (Bleidorn & Schwaba, 2018; Schwaba & Bleidorn, 2019). These results are only partly in line with the social investment principle regarding mean-level changes and display substantial interindividual differences in change trajectories. Schwaba and Bleidorn described retirement as a “divestment” of social roles (2019, p. 660; for personality relaxation, see Asselmann & Specht, 2021) that functions differently than social investment, which adds a role. The grandparent role is one of only a few new normative roles available in middle and older adulthood. It is perceived as highly important and represents a psychologically meaningful role investment (Mahne & Motel-Klingebiel, 2012; Thiele & Whelan, 2006)—given that grandparents have regular contact with grandchildren and take part in childcare (Lodi-Smith & Roberts, 2007). Mechanisms of grandparents’ personality change remain unexplored. However, grandparental role investment may not be linearly related to changes in well-being and health (see section Life satisfaction and grandparenthood). Instead, moderate levels of grandchild care and contact appear most beneficial. At the same time, even if grandparents do not provide substantial or regular grandchild care, the transition to grandparenthood might still alter their everyday lives and activities considerably by changing the social structure imposed by kinship bonds (Mueller & Elder, 2003; Tanskanen, 2017). For example, grandchildren might bring about frequent family gatherings, which eventually contribute to grandparents’ personality development in a bottom-up fashion.
Grandparenthood

The transition to grandparenthood is a time-discrete life event—the beginning of one’s status as a grandparent (Luhmann et al., 2012). In terms of characteristics of major life events (Luhmann et al., 2020), the transition to grandparenthood stands out in that it is externally caused (by one’s children; see also Arpino, Gumà, et al., 2018; Margolis & Verdery, 2019), but also predictable as soon as children reveal their family planning or pregnancy. The transition to grandparenthood has been labeled a counter-transition due to this lack of direct control over its timing (Hagedoost & Neugarten, 1985; as cited in Arpino, Gumà, et al., 2018). Grandparenthood is also generally positive in valence and emotionally significant if the grandparent maintains a good relationship with their child. Grandparents’ investments in their grandchildren are beneficial in terms of the evolutionary, economic, and sociological advantages they provide (Coall et al., 2018; Coall & Hertwig, 2011).

Grandparenthood is a developmental task (Hutteman et al., 2014) that generally takes place in (early) old age, although this varies considerably both within and between cultures (Leopold & Skopek, 2015; Skopek & Leopold, 2017). Still, the period in which parents experience the birth of their first grandchild coincides with the end of (relative) personality stability in midlife (Specht, 2017), when retirement, shifting social roles, and initial cognitive and health declines can disrupt life circumstances, setting processes of personality development in motion (e.g., Mueller et al., 2016; Stephan et al., 2014). As a developmental task, grandparenthood is considered part of a normative sequence of aging that is subject to societal expectations and values that differ across cultures and historical time (Baltes et al., 2006; Hutteman et al., 2014). Mastering developmental tasks (i.e., fulfilling roles and expectations) is hypothesized to drive positive personality development similarly to propositions of the social investment principle, that is, leading to higher levels of agreeableness and conscientiousness, and lower levels of neuroticism (Roberts et al., 2005; Roberts & Wood, 2006).

In comparison to the transition to parenthood, which is ambivalent in terms of both personality maturation and changes in life satisfaction (Aasve et al., 2021; Johnson & Rodgers, 2006; Krämer & Rodgers, 2020; van Scheppingen et al., 2016), Hutteman et al. (2014) hypothesized that the transition to grandparenthood is positive because it (usually) does not impose the stressful demands of daily childcare on grandparents. However, societal expectations about how grandparents should behave are less clearly defined than expectations around parenthood. There is considerable heterogeneity in how intensely grandparents are involved in their grandchildren’s lives and care (Meyer & Kandic, 2017). The degree of possible grandparental investment differs depending on a variety of factors: how close grandparents live to their children, the quality of their relationship, and sociodemographic factors that create conflicting role demands such as paid work or other caregiving responsibilities (Arpino & Bellani, 2022; Arpino & Gómez-León, 2020; Lumsdaine & Vermeer, 2015; Silverstein & Marenco, 2001). In the entire population of first-time grandparents, this diversity of possible and desired role investments could generate role conflicts for some grandparents (according to role strain theory; Goode, 1960). Subsequently, pronounced interindividual differences in intraindividual personality change might then emerge.

Life satisfaction and grandparenthood. Although few studies on the Big Five and grandparenthood exist, there is some evidence for life satisfaction, which we define as the general, cognitive appraisal of one’s well-being in life based on subjective criteria (Eid & Larsen, 2008). Life satisfaction is generally considered less stable than the Big Five and more prone to changes due to environmental influences but still trait-like in its characteristics (Anusic & Schimmack, 2016; Kandler et al., 2014; Luhmann et al., 2012), and robustly related to the Big Five (Anglim et al., 2020).

Longitudinal studies on grandparents’ life satisfaction have produced conflicting conclusions: Studies using data from the Survey of Health, Aging and Retirement in Europe (SHARE) showed that the birth of a grandchild was followed by improvements in quality of life and life satisfaction, but only among women (Tanskanen et al., 2019) and only in first-time grandmothers via their daughters (Di Gessa et al., 2019). Several studies demonstrated that grandparents who were actively involved in childcare experienced larger increases in life satisfaction (Arpino, Bordone, et al., 2018; Danielsbacka et al., 2019; Danielsbacka & Tanskanen, 2016). On the other hand, fixed effects regression models using SHARE data did not find any effects of first-time grandparenthood on life satisfaction regardless of grandparental investment and only minor decreases in depressive symptoms in grandmothers (Sheppard & Monden, 2019; see also Ates, 2017, who came to a similar conclusion for self-rated health using data from the German Aging Survey).

Studies of grandparents’ life satisfaction, and well-being and health more generally, have often contrasted role strain theory and role enhancement theory (e.g., Di Gessa et al., 2016a; Xu et al., 2017; see also Kim et al., 2017). Role strain theory (Goode, 1960) predicts that investing in grandparenthood alongside other existing roles can produce role conflicts and psychological demands exceeding one’s resources. Altogether, these factors prevent adaptive development and lower life satisfaction. Role enhancement theory (Sieber, 1974), conversely, anticipates adaptive development and well-being benefits because the added social role provides grandparents with status security, social support, and psychological meaning. Empirically, providing grandchild care is, on the one hand, associated with decreased marital satisfaction (Wang & Mitchler, 2020) and increased depressive symptoms if grandparents perceive caregiving as burdensome (Xu et al., 2017). On the other hand, it is associated with increased social contact (Quirke et al., 2021; Tanskanen, 2017; cf. Arpino & Bordone, 2017) and a higher quantity (but not quality) of leisure activities (Ates et al., 2021), whereby social engagement serves as a buffer for mental health decreases (Notter, 2022).

Research on well-being and health has found evidence for both role strain theory and role enhancement theory depending on the degree of grandparental role investment (Danielsbacka et al., 2022; Kim et al., 2017). Whereas no
investment or being a grandchild’s primary caregiver are associated with adverse effects in most studies, there is evidence that moderate levels of grandchild care have beneficial life satisfaction and health effects for noncoresiding grandparents. This provides preliminary support for the inverted U-shape between investment and utility proposed by Coall and Hertwig (2011). However, multiple authors have recently emphasized that the literature is still at an early stage and that prior studies often lack representativeness, longitudinal data, and appropriate control for selection effects (Coall et al., 2018; Danielsbacka et al., 2022; Kim et al., 2017).

In summary, evidence is lacking on the Big Five and inconclusive on life satisfaction (and related measures) which is partly due to different methodological approaches that do not account for confounding (i.e., selection effects).

**Methodological considerations**

Effects of life events on psychological traits tend to be small and need to be analyzed using robust, prospective designs and appropriate control groups (Bleidorn et al., 2018; Luhmann et al., 2014). This is necessary because pre-existing differences between prospective grandparents and non-grandparents in variables related to the development of the Big Five or life satisfaction introduce confounding bias when estimating the effects of the transition to grandparenthood (VanderWeele et al., 2020). The impact of adjusting for pre-existing differences was recently emphasized in predicting life outcomes from personality (Beck & Jackson, 2022). Propensity score matching is one technique to account for confounding bias by equating groups in their estimated propensity to experience the event (Thoemmes & Kim, 2011). This propensity is calculated from regressing the so-called treatment variable (whether someone experienced the event) on covariates related to the likelihood of experiencing the event and to the outcomes. This approach addresses confounding bias by creating balance between groups in the covariates used to calculate the propensity score (Stuart, 2010).

We adopt a prospective design that tests the effects of becoming first-time grandparents against two propensity-score-matched control groups separately: first, parents (but not grandparents) with at least one child, and, second, nonparents. This allows us to disentangle potential effects of becoming a grandparent from effects of already being a parent (i.e., parents who eventually become grandparents might share additional similarities with parents who do not). Thus, we can address selection effects into grandparenthood more comprehensively than previous research. We cover the first two of three causal pathways to not experiencing grandparenthood pointed out in demographic research (Margolis & Verdery, 2019): childlessness, childlessness of one’s children, and not living long enough to become a grandparent. Our comparative design controls for average age-related and historical trends in the Big Five traits and life satisfaction (Luhmann et al., 2014). The design also enables us to report effects of the transition to grandparenthood unconfounded by instrumentation effects, which describe the tendency of reporting lower well-being scores with each repeated measurement (Baird et al., 2010).

We match at a specific time point before the transition to grandparenthood (i.e., at least 2 years beforehand) and not based on individual survey years. This design choice ensures that the covariates involved in the matching procedure are not already influenced by the event or anticipation of it (Greenland, 2003; Rosenbaum, 1984; VanderWeele, 2019; VanderWeele et al., 2020), thereby reducing the risk of introducing confounding through collider bias (Elwert & Winship, 2014). Similar approaches in the study of life events have been adopted recently (Balbo & Arpino, 2016; Krämer & Rodgers, 2020; van Scheppingen & Leopold, 2020).

**Current study**

In the current study, we examine the development of the Big Five personality traits across the transition to grandparenthood in a prospective, quasi-experimental design, thereby extending previous research on the effects of this transition on well-being to psychological development in a more general sense. We also revisit life satisfaction development, which allows us to anchor our model results. With the literature on grandparenthood and well-being in mind, the current results for life satisfaction constitute a benchmark for the Big Five outcomes. Three research questions motivate the current study which—to our knowledge—is the first to analyze Big Five personality development over the transition to grandparenthood:

1. What are the effects of the transition to grandparenthood on mean-level trajectories of the Big Five traits and life satisfaction?
2. How large are interindividual differences in intra-individual change for the Big Five traits and life satisfaction over the transition to grandparenthood?
3. How does the transition to grandparenthood affect rank-order stability of the Big Five traits and life satisfaction?

To address these questions, we used two nationally representative panel data sets and compared grandparents’ development over the transition to grandparenthood with that of matched respondents who did not become grandparents during the study period (Luhmann et al., 2014). Informed by the social investment principle, previous research on personality development in middle and older adulthood, and the literature on grandparenthood and well-being, we preregistered the following hypotheses (see https://osf.io/a9zpc):

- **H1a:** Following the birth of their first grandchild, grandparents increase in agreeableness and conscientiousness, and decrease in neuroticism compared to the matched control groups of parents (but not grandparents) and nonparents. We do not expect the groups to differ in their trajectories of extraversion and openness to experience.
- **H1b:** Grandparents’ post-transition increases in agreeableness and conscientiousness, and decreases in neuroticism are more pronounced among those who provide substantial grandchild care.
The heterogeneity in the degree of possible and desired grandparental investment in our samples leads us to expect pronounced interindividual differences in intraindividual change (i.e., deviations from the average trajectories).

- **H1c**: Grandmothers increase in life satisfaction following the transition to grandparenthood compared to the matched control groups but grandfathers do not.

Consequently, assuming that grandparents’ personality is rearranged through the experience of the event, we also expect decreases in rank-order stability over the transition to grandparenthood.

- **H2**: Individual differences in intraindividual change in the Big Five and life satisfaction are larger in the grandparent group than the control groups.

Finally, commitments to other institutions and roles possibly constrain the amount of possible grandparental investment in line with role strain theory. Alternatively, the added grandparental role could complement existing roles inducing positive psychological development according to role enhancement theory. Thus, exploratorily, we probe the added grandparental role could complement existing roles.

- **H3**: Compared to the matched control groups, grandparents’ rank-order stability of the Big Five and life satisfaction over the transition to grandparenthood is smaller.

**Methods**

**Samples**

We used data from two population-representative panel studies: The Longitudinal Internet Studies for the Social Sciences (LISS) panel from the Netherlands, and the Health and Retirement Study (HRS) from the United States.

The LISS panel is a representative sample of the Dutch population initiated in 2008 with data collection still ongoing (Scherpenzeel, 2011; van der Laan, 2009). It is administered by Center data (Tilburg University). The survey population is a true probability sample of households drawn from the population register (Scherpenzeel & Das, 2010). Data collection was carried out online, and respondents were provided technical equipment if needed. We included yearly assessments from 2008 to 2021 as well as basic demographics assessed monthly. For later coding of covariates from these monthly demographic data we used the first available assessment each year.

The HRS is an ongoing population-representative study of older adults in the United States (Sonnega et al., 2014) administered by the Survey Research Center (University of Michigan). Initiated in 1992 with a first cohort of individuals aged 51–61 and their spouses, the study has since been expanded through additional cohorts (see https://hrs.isr.umich.edu/documentation/survey-design/). In addition to the biennial in-person or telephone interview, since 2006 the study has included a leave-behind questionnaire covering psychosocial topics including personality traits. These topics, however, were only administered every 4 years starting in 2006 for one half of the sample and in 2008 for the other half. We included personality data from 2006 to 2018, all available data for the coding of the transition to grandparenthood from 1996 to 2018, as well as covariate data from 2006 to 2018 including variables drawn from the Imputations File and the Family Data (only available up to 2014).

These two panel studies provided the advantage that they contained several waves of personality data as well as information on grandparent status and a broad range of covariates. While the HRS provided a large sample with a wider age range, the LISS was smaller and younger but provided more frequent personality assessments spaced every 1–2 years. Included grandparents from the LISS were younger because grandparenthood questions were part of the Work and Schooling module and—for reasons unknown to us—filtered to respondents performing paid work. Thus, older, retired first-time grandparents from the LISS could not be identified. Even though we have published using the LISS and HRS data before (see https://osf.io/a9zpc), these publications do not overlap with the current study on grandparenthood. The present study used de-identified archival data available in the public domain, which meant that it was not necessary to obtain ethical approval from an IRB.

**Measures**

**Personality.** In the LISS, the Big Five personality traits were assessed using the 50-item version of the IPIP Big Five Inventory scales (Goldberg, 1992). For each trait, respondents answered ten 5-point Likert-scale items (1 = very inaccurate, 2 = moderately inaccurate, 3 = neither inaccurate nor accurate, 4 = moderately accurate, 5 = very accurate). Example items included “like order” (conscientiousness), “sympathize with others’ feelings” (agreeableness), “worry about things” (neuroticism), “have a vivid imagination” (openness), and “start conversations” (extraversion). In each wave, we took a respondent’s mean of each subscale as their trait score. Internal consistencies at the time of matching, as indicated by $\omega_h$ (McNeish, 2018), averaged $\omega_h = 0.70$ over all traits ($\omega_i = 0.89; \alpha = 0.83$; see Table S1). Other studies have shown measurement invariance for these scales across time and age groups, and convergent validity with the Big Five Inventory (BFI-2; Schwaba & Bleidorn, 2018; Denissen et al., 2020). The Big Five and life satisfaction were administered yearly but with planned missingness in some years for certain cohorts (see Denissen et al., 2019).

In the HRS, the Midlife Development Inventory (MIDI) scales measured the Big Five (Lachman & Weaver, 1997) with 26 adjectives (five for conscientiousness, agreeableness, and extraversion; four for neuroticism; seven for openness). Respondents were asked to rate on a 4-point scale how well each item described them ($1 = a lot$, $2 = some$, $3 = a little$, $4 = not at all$). Example adjectives included...
organized (conscientiousness), sympathetic (agreeableness), worrying (neuroticism), imaginative (openness), and talkative (extraversion). For better comparability with the LISS panel, we reverse-scored all items so that higher values corresponded to higher trait levels and, in each wave, took the mean of each subscale as the trait score. Big Five trait scores showed satisfactory internal consistencies at the time of matching that averaged $\omega_h = 0.63$ over all traits ($\omega_t = 0.80; \alpha = 0.72$; see Table S1).

**Life satisfaction.** In both samples, life satisfaction was assessed using the 5-item Satisfaction with Life Scale (SWLS; Diener et al., 1985) which respondents answered on a 7-point Likert scale (1 = strongly disagree, 2 = somewhat disagree, 3 = slightly disagree, 4 = neither agree or disagree, 5 = slightly agree, 6 = somewhat agree, 7 = strongly agree). An example item was “I am satisfied with my life”. Internal consistency at the time of matching was between $\alpha = 0.80$ and $\alpha = 0.91$ in the four analysis samples (see Table S1).

**Transition to grandparenthood.** The procedure to obtain information on the transition to grandparenthood generally followed the same steps in both samples. This coding was based on items that differed slightly, however: In the LISS, respondents performing paid work were asked “Do you have children and/or grandchildren?” and were offered the answer categories “children,” “grandchildren,” and “no children or grandchildren.” In the HRS, all respondents were asked to state their total number of grandchildren: “Altogether, how many grandchildren do you (or your husband / wife / partner, or your late husband / wife / partner) have? Include as grandchildren any children of your (or your [late] husband’s/wife’s / partner’s) biological, step- or adopted children.”

In both samples, we tracked grandparenthood status over time using all available longitudinal information (including HRS waves 1996–2018). Due to longitudinally inconsistent data in some cases, we included in the grandparent group only respondents with one transition from 0 (no grandchildren) to 1 (at least one grandchild) in this status variable, and no transitions backwards (see Figure 1). We marked respondents who consistently indicated that they had no grandchildren as potential members of the control groups.

**Moderators.** We tested four variables as potential moderators of the mean-level trajectories of the Big Five and life satisfaction over the transition to grandparenthood: First, we analyzed whether female gender (0 = male, 1 = female) acted as a moderator as indicated by research on life satisfaction (Di Gessa et al., 2019; Tanskanen et al., 2019).

Second, we tested whether performing paid work (0 = no, 1 = yes) was associated with divergent trajectories of the Big Five and life satisfaction (Schwaba & Bleidorn, 2019). Since the LISS subsample consisted solely of respondents...
performing paid work, we performed these analyses only in the HRS. This served two purposes. On the one hand, it allowed us to test how respondents in the workforce differed from those not working, which might shed light on role conflict and have implications for social investment mechanisms. On the other hand, these moderation analyses allowed us to assess whether potential differences in results between the LISS and HRS samples could be accounted for by including performing paid work as a moderator in HRS analyses. In other words, perhaps HRS respondents performing paid work were similar to those in the LISS sample—those conditioned on this variable through questionnaire filtering.

Third, we examined how involvement in grandchild care moderated trajectories of the Big Five and life satisfaction (Arpino, Bordone, et al., 2018; Danielsbacka et al., 2019; Danielsbacka & Tanskanen, 2016). We coded a moderator variable (0 = provided less than 100 hours of grandchild care, 1 = provided 100 or more hours of grandchild care) based on the question “Did you (or your [late] husband / wife / partner) spend 100 or more hours in total since the last interview / in the last two years taking care of grand- or great grandchildren?” This information was only available for grandparents in the HRS (43% yes); in the LISS, too few respondents answered respective follow-up questions to be included in analyses.

Fourth, in the HRS, we compared Black/African American respondents with White respondents.

**Procedure**

Drawing on all available data, three main restrictions defined the analysis samples of grandparents (see Figure 1): First, we identified respondents who indicated having grandchildren for the first time during study participation ($N_{LISS} = 380; N_{HRS} = 3273$, including HRS waves 1996–2004 before personality assessments were introduced). Second, we restricted the sample to respondents with at least one valid personality assessment (valid in the sense that at least one of the six outcomes was non-missing; $N_{LISS} = 378; N_{HRS} = 1703$). Third, we included only respondents with both one valid personality assessment before and one after the transition to grandparenthood ($N_{LISS} = 283; N_{HRS} = 860$). Finally, a few respondents were excluded because of inconsistent or missing information regarding their children resulting in the final analysis samples of first-time grandparents, $N_{LISS} = 282$ (54.61% female; age at transition to grandparenthood $M = 58.29$, $SD = 4.87$) and $N_{HRS} = 847$ (54.90% female; age at transition to grandparenthood $M = 61.80$, $SD = 6.87$).

We defined two mutually exclusive pools of potential control subjects for matching: The first comprised parents who had at least one child (given that $15 \leq \text{age of youngest} \leq 65$) but no grandchildren during the observation period ($N_{LISS} = 853$ with 3846 longitudinal observations; $N_{HRS} = 1485$ with 2703 longitudinal observations). The second comprised respondents who reported being childless throughout the observation period ($N_{LISS} = 986$ with 4906 longitudinal observations; $N_{HRS} = 1340$ with 2346 longitudinal observations).

**Covariates.** We used propensity score matching to match each grandparent with a control respondent from each pool of potential controls who was most similar in terms of the included covariates.

Although critical to the design, covariate selection is seldom explicitly discussed in studies estimating effects of life events (e.g., in matching designs). We see two (in part conflicting) traditions that address covariate selection: First, classic recommendations from psychology are to include all available variables that are associated with both the treatment assignment process (i.e., selection into treatment) and the outcome (e.g., Steiner et al., 2010; Stuart, 2010). Second, recommendations from a structural causal modeling perspective (Elwert & Winship, 2014; Rohrer, 2018) are more cautious, aiming to avoid pitfalls such as conditioning on a pre-treatment collider (collider bias) or a mediator (overcontrol bias). However, structural causal modeling requires advanced knowledge of the causal structures underlying the involved variables (Pearl, 2009).

In selecting covariates, we followed the guidelines of VanderWeele (2019) and VanderWeele et al. (2020), which reconcile both views and offer practical guidance when the underlying causal structures are not completely understood and when using large archival datasets. The “modified disjunctive cause criterion” (VanderWeele, 2019, p. 218) recommends selecting all available covariates which are assumed to be causes of the outcomes, treatment exposure (i.e., the transition to grandparenthood), or both, as well as any proxies for an unmeasured common cause of the outcomes and treatment exposure. Variables that are assumed to be instrumental variables (i.e., assumed causes of treatment exposure that are unrelated to the outcomes except through the exposure) and collider variables (Elwert & Winship, 2014) should be excluded from this selection.

Because all covariates we used for matching were measured at least 2 years before the first grandchild’s birth, we judge the risk of introducing collider bias or overcontrol bias to be relatively small. In addition, as mentioned above, the transition to grandparenthood is not planned by or under the direct control of the grandparents, which further reduces the risk of these biases.

Following these guidelines, we selected covariates covering respondents’ demographics (e.g., age, education), economic situation (e.g., income), and health (e.g., mobility difficulties). We also included the pre-transition outcome variables as covariates—as recommended in the literature (Cook et al., 2020; Hallberg et al., 2018; Steiner et al., 2010; VanderWeele et al., 2020), as well as wave participation count and assessment year in order to control for instrumentation effects and historical trends (e.g., 2008/2009 financial crisis; Baird et al., 2010; Luhmann et al., 2014). To match grandparents with the parent control group, we additionally selected covariates containing information on fertility and family history (e.g., number of children, age of first three children) which were causally related to the timing of the transition to grandparenthood (Arpino, Gumà, et al., 2018; Margolis & Verdery, 2019).

An overview of all covariates can be found in the supplemental materials (see Tables S2 & S3). Importantly, as part of our preregistration we justified each covariate, explaining whether we assumed it to be related to the treatment assignment, the outcomes, or both (see gp-covariates-overview.xlsx on https://osf.io/75a4r/). In this
document, we provided references supporting our assumptions on whether a specific covariate is related to these causal processes. For example, we justified the inclusion of religion as a covariate with its relation to fertility (Hayford & Morgan, 2008; Zhang, 2008), which is often passed down to the child’s family (Götmark & Andersson, 2020), and its relation to the Big Five and life satisfaction (Diener et al., 2018; Gebauer et al., 2014). We tried to find substantively equivalent covariates in both samples but had to compromise in a few cases.

Estimating propensity scores required complete covariate data. Therefore, we performed multiple imputations to address missingness in the covariates (Greenland & Finkle, 1995). Using five imputed data sets computed by classification and regression trees (CART; Burgette & Reiter, 2010) in the mice R package (van Buuren & Groothuis-Oudshoorn, 2011), we predicted treatment assignment (i.e., the transition to grandparenthood) five times per observation in logistic regressions with a logit link function. We averaged these five scores per observation to compute the final propensity score used for matching (Mitra & Reiter, 2016). We used imputed data only for propensity score computation and not in later analyses because nonresponse in the outcome variables was negligible.

**Propensity score matching.** The time of matching preceded the survey year in which the transition to grandparenthood was first reported by at least 2 years (aside from that choosing the smallest available gap between matching and transition). This ensured that the covariates were not affected by the event itself or anticipation thereof (i.e., matching occurred well before children would have announced that they were expecting their first child; Greenland, 2003; Rosenbaum, 1984; VanderWeele et al., 2020). Propensity score matching was performed using the Matchit R package (Ho et al., 2011) with exact matching on gender combined with Mahalanobis distance matching on the propensity score. Four matchings were performed; two per sample (LISS; HRS) and two per control group (parents; nonparents). We matched 1:4 with replacement because of the relatively small pools of available controls. We did not specify a caliper because our goal was to find matches for all grandparents, and because we achieved good covariate balance this way.

We evaluated the matching procedure in terms of covariate balance and graphically (Stuart, 2010). Covariate balance as indicated by the standardized difference in means between grandparents and controls after matching was good (see Tables S2 & S3), lying below 0.25 as recommended in the literature (Stuart, 2010), and below 0.10 with few exceptions (Austin, 2011). Graphically, group differences in the propensity score distributions were small and indicated no substantial missing overlap (see Figure S1).

After matching, each matched control observation was assigned the same value as the matched grandparent in the time variable describing the temporal relation to treatment, and the control respondent’s other longitudinal observations were centered around this matched observation. We thus coded a counterfactual transition time frame for each control respondent. Due to left- and right-censored longitudinal data (i.e., panel entry or attrition), we restricted the final analysis samples to 6 years before and 6 years after the transition, as shown in Table 1.

The final LISS analysis samples (see Figure 1) contained 282 grandparents with 1591 longitudinal observations, matched with 1128 control respondents with either 6288 (parent control group) or 6290 longitudinal observations (nonparent control group). The final HRS analysis samples contained 847 grandparents with 2264 longitudinal observations, matched with 3388 control respondents with either 8326 (parent control group) or 8229 longitudinal observations (nonparent control group). In the HRS, there were a few additional missing values in the outcomes ranging from 19 to 99 longitudinal observations, which were listwise deleted in the respective analyses.

**Transparency and openness**

We used R (Version 4.0.4; R Core Team, 2021) and the R-packages lme4 (Version 1.1.27.1; Bates et al., 2015), and lmerTest (Version 3.1.3; Kuznetsova et al., 2017) for multilevel modeling, as well as tidyverse (Wickham et al., 2019) for data wrangling, and papaja (Aust & Barth, 2020) for reproducible manuscript production (see supplement for complete package information). The preregistration and scripts for data wrangling, analyses, and to reproduce this manuscript can be found on the OSF (https://osf.io/75a4r/) and GitHub (https://github.com/mdkraemer/gp-personality). LISS and HRS data are available after registering accounts. We deviate from the preregistration in using new waves of data released in the meantime (2020/2021 LISS) as well as updated datasets (HRS). Following Benjamin et al. (2018), we set the α-level for confirmatory analyses to 0.005.

**Analytical strategy**

Our design can be referred to as an interrupted time series with a “nonequivalent no-treatment control group” (Shadish et al., 2002, p. 182) where treatment, that is, the transition to grandparenthood, is not deliberately manipulated. First, to analyze mean-level changes (research question 1), we used linear piecewise regression coefficients: a before-slope representing linear change in the years leading up to the transition to grandparenthood, an after-slope representing linear change in the years after the transition, and a shift coefficient, shifting the intercept directly after the transition was first reported, thus representing sudden changes that go beyond changes already modeled by the after-slope (see Table 1 for the coding scheme of these coefficients).

Other studies of personality development have recently adopted similar piecewise coefficients (e.g., Krämer & Rodgers, 2020; Schwaba & Bleidorn, 2019; van Scheppingen & Leopold, 2020).
All effects of the transition to grandparenthood on the Big Five and life satisfaction were modeled as deviations from the matched control groups by interacting the three piecewise coefficients with the treatment variable (0 = control, 1 = grandparent). In additional models, we interacted these coefficients with the moderator variables, resulting in two- and three-way interactions. To test differences in the growth parameters between two groups in cases where these differences were represented by multiple fixed-effects coefficients, we defined linear contrasts using the `linearHypothesis` command from the `car` package (Fox & Weisberg, 2019). All models of mean-level changes were estimated using maximum likelihood and included random intercepts but no random slopes. Simultaneous random slopes of change parameters frequently lead to convergence issues. Fixed slopes models are appropriate to model average trajectories, which vary systematically with the person-level treatment variable (Hoffman & Walters, 2022). We included the propensity score as a level-2 covariate for a double-robust approach (Austin, 2021). The equation for the basic (i.e., unmoderated) model reads

\[ y_{it} = \beta_0 + \beta_1, before_{it} + \beta_2, after_{it} + \beta_3, shift_{it} + e_{it} \]

\[ \beta_0 = \gamma_{00} + \gamma_{01}, grandparent, + \gamma_{02}, pscore, + \upsilon_{0i} \]

\[ \beta_1 = \gamma_{10} + \gamma_{11}, grandparent, + \upsilon_{1i} \]

\[ \beta_2 = \gamma_{20} + \gamma_{21}, grandparent, \]

\[ \beta_3 = \gamma_{30} + \gamma_{31}, grandparent, \]

where at time \( t \) for person \( i \), \( e_{it} \) is \( \sim N(0, \sigma^2) \) and \( \upsilon_{0i} \) is \( \sim N(0, \tau_{00}) \) (ignoring the additional nesting in households applied to the majority of models). \( y_{it} \) represented one of the Big Five or life satisfaction. Separate models were computed for each analysis sample. The other model equations can be found in the supplemental materials.

Second, to assess interindividual differences in change (research question 2), we added random slopes. In other words, we allowed for differences between individuals in their trajectories of change to be modeled, that is, differences in the \( before-slope, after-slope, \) and \( shift \) coefficients. Because simultaneous random slopes are often not computationally feasible, we added random slopes one at a time and used likelihood ratio tests to determine whether the addition of the respective random slope led to a significant improvement in model fit. To test differences in the random slope variance between the grandparent group and each control group, we respecified the models as heterogeneous variance models using the `nlme` R package (Pinheiro et al., 2021). This allowed for separate random slope variances to be estimated in the grandparent group and the control group within the same model. We compared the fit of these heterogeneous variance models to corresponding models with a homogeneous (single) random slope variance using likelihood ratio tests.

Third, to examine rank-order stability in the Big Five and life satisfaction over the transition to grandparenthood (research question 3), we computed the test-retest correlation of measurements prior to the transition to grandparenthood (at the time of matching) and the first available measurement afterward. To test differences in test-retest correlations between grandparents and either of the control groups, we entered the pre-treatment measure, the treatment variable (0 = control, 1 = grandparent), and their interaction into regression models predicting the Big Five and life satisfaction.
satisfaction. The interaction tests for significant differences in the rank-order stability between those who experienced the transition to grandparenthood and those who did not (see Denissen et al., 2019; McCrae, 1993).

**Results**

Throughout the results section, we referred to statistical tests with \( p < .05 \) as suggestive evidence as stated in our preregistration.

**Descriptive results**

Means and standard deviations of the Big Five and life satisfaction over the analyzed time points are presented in Tables S4 and S5. Visually represented (see Figures S2–S7), all six outcomes display marked stability over time in both LISS and HRS. Intra-class correlations (see Table S6) show that large portions of the total variance in the Big Five could be explained by nesting in respondents (median = 0.75), while nesting in households only accounted for minor portions of the total variance (ICC\textsubscript{hid}, median = 0.03). For outcome-subsample combinations with ICC\textsubscript{hid} below 0.05 we omitted the household nesting factor from all models to bypass computational errors—a small deviation from our preregistration. For life satisfaction, the nesting in households accounted for slightly larger portions of the total variance (median = 0.37) than nesting in respondents (median = 0.30). Across all outcomes, the proportion of variance due to within-person factors was relatively low (median = 0.23).

**Mean-level changes**

Figures 2 and 3 summarize the effects of the basic models and those including the gender interaction for all outcomes and across the four analysis samples.

**Agreeableness.** In the basic models, we found no evidence that grandparents increased in agreeableness as compared to the controls (see Tables S7 & S8 and Figure 4). The models including the gender interaction (Tables 2 & S9 and Figure 4) indicated that grandfathers increased slightly in agreeableness after the transition to grandparenthood as compared to the parent controls (LISS: \( \gamma_{21} = 0.02, 95\% \text{ CI} [0.01, 0.04], p = .002 \); suggestive evidence in the HRS: \( \gamma_{21} = 0.03, 95\% \text{ CI} [0.01, 0.05], p = .008 \)), whereas grandmothers did not differ from the female controls. There was no consistent evidence for moderation by paid work (see Tables S10 & S11 and Figure S8), providing grandchild care (see Tables S12 & S13 and Figure S9), or ethnicity (see Tables S14 & S15 and Figure S10).

**Conscientiousness.** We found no differences between grandparents and both parent and nonparent controls in their trajectories of conscientiousness (see Tables S16 & S17 and Figure S11). There was only inconsistent evidence for gender moderation (see Tables S18 & S19 and Figure S11): Grandfathers’ conscientiousness decreased immediately following the transition to grandparenthood as compared to male nonparents in the HRS, \( \gamma_{21} + \gamma_{31} = -0.07, 95\% \text{ CI} [-0.11, -0.02], p = .004 \), but not in any of the other three analysis samples.

There were significant differences in conscientiousness trajectories depending on grandparents’ work status (see Tables 3 & S20 and Figure 5): non-working grandparents saw more pronounced increases in conscientiousness in the years before the transition to grandparenthood compared to non-working parents, \( \gamma_{21} = 0.08, 95\% \text{ CI} [0.03, 0.13], p < .001 \), and nonparent controls, \( \gamma_{31} = 0.06, 95\% \text{ CI} [0.02, 0.11], p = .004 \), and compared to working grandparents (difference in \textit{before} parameter; parents: \( \gamma_{20} + \gamma_{31} = -0.08, 95\% \text{ CI} [-0.13, -0.03], p = .002 \); nonparents: \( \gamma_{20} + \gamma_{31} = -0.08, 95\% \text{ CI} [-0.12, -0.03], p = .001 \)). Grandparents providing grandchild care increased in conscientiousness to a greater degree than the matched controls (difference in \textit{after} parameter; parents: \( \gamma_{21} + \gamma_{31} = 0.04, 95\% \text{ CI} [0.02, 0.06], p < .001 \); nonparents: \( \gamma_{21} + \gamma_{31} = 0.04, 95\% \text{ CI} [0.02, 0.06], p < .001 \); see Tables 4 & S21 and Figure 6). There was only suggestive evidence that grandparents who provided grandchild care increased more strongly in conscientiousness after the transition than grandparents who did not (difference in \textit{after} parameter; parents: \( \gamma_{20} + \gamma_{31} = 0.03, 95\% \text{ CI} [0.00, 0.06], p = .029 \); nonparents: \( \gamma_{20} + \gamma_{31} = 0.03, 95\% \text{ CI} [0.01, 0.06], p = .020 \)). Conscientiousness trajectories were not moderated by ethnicity (see Tables S22 & S23 and Figure S12).

**Extraversion.** The trajectories of grandparents’ extraversion closely followed those of the matched controls. There were no significant effects indicating differences between grandparents and controls in the basic models (see Tables S24 & S25 and Figure S13) or the models including the gender interaction (see Tables S26 & S27 and Figure S13). We also found no evidence for moderation by paid work (see Tables S28 & S29 and Figure S14), grandchild care (see Tables S30 & S31 and Figure S15), or ethnicity (see Tables S32 & S33 and Figure S16).

**Neuroticism.** The basic models for neuroticism (see Tables S34 & S35 and Figure S17) showed only minor differences between grandparents and matched controls: Compared to HRS parent controls, HRS grandparents shifted slightly downward in their neuroticism immediately after the transition to grandparenthood (difference in \textit{shift} parameter: \( \gamma_{21} + \gamma_{31} = -0.07, 95\% \text{ CI} [-0.11, -0.02], p = .003 \); suggestive evidence in the nonparent sample: \( \gamma_{21} + \gamma_{31} = -0.05, 95\% \text{ CI} [-0.09, 0.00], p = .042 \), which was not the case in the LISS samples. The models including the gender interaction (see Tables S36 & S37 and Figure S17) showed one significant effect in the comparison of grandparents and controls: In the HRS, grandfathers, compared to male parent controls, shifted downward in neuroticism directly after the transition to grandparenthood (difference in \textit{shift} parameter: \( \gamma_{21} + \gamma_{31} = -0.15, 95\% \text{ CI} [-0.21, -0.08], p < .001 \)). Thus, the effect present in the basic models seemed to be mostly due to differences in the grandfathers (vs. male controls).

Grandparents’ trajectories of neuroticism as compared to the controls were significantly moderated by paid work in one instance (see Tables S38 & S39 and Figure S18): Compared to working controls, working grandparents...
Figure 2. Unstandardized effect sizes of the basic models across analysis samples. Depicted are regression coefficients $\hat{\gamma}$ or linear contrasts $\hat{\gamma}_c$ from multilevel models, see Tables S7, S8, S16, S17, S24, S25, S34, S35, S44, S45, S54, S55. Error bars represent 95% confidence intervals.

Figure 3. Unstandardized effect sizes of the models including the gender interaction across analysis samples. Depicted are regression coefficients $\hat{\gamma}$ or linear contrasts $\hat{\gamma}_c$ from multilevel models, see Tables 2, S9, S18, S19, S26, S27, S36, S37, S46, S47, S56, S57. Error bars represent 95% confidence intervals.
increased more strongly in neuroticism in the years before the transition to grandparenthood (difference in *before* parameter; parents: $\hat{\beta}_{31} + \hat{\gamma}_{31} = 0.06$, 95% CI [0.02, 0.10], $p = .001$; nonparents: $\hat{\beta}_{31} + \hat{\gamma}_{31} = 0.06$, 95% CI [0.02, 0.09], $p = .002$). There was no evidence that grandparents providing grandchild care differed in neuroticism from grandparents who did not (see Tables S40 & S41 and Figure S19). Neuroticism trajectories were not moderated by ethnicity (see Tables S42 & S43 and Figure S20).

**Openness.** For openness, we found a high degree of similarity between grandparents and matched control respondents in their trajectories based on the basic models (see Tables S44 & S45 and Figure S21) and models including the gender interaction (see Tables S46 & S47 and Figure S21). Grandfathers in the HRS shifted downward in openness in the first assessment after the transition to grandparenthood to a greater extent than the male parent controls (difference in *shift* parameter: $\hat{\beta}_{31} + \hat{\gamma}_{31} = -0.09$, 95% CI [-0.14, -0.03], $p = .002$). However, this was not the case in the other three analysis samples.

The analysis of moderation by performing paid work revealed only one significant effect for openness trajectories (see Tables S48 & S49 and Figure S22): Non-working grandparents increased more strongly in openness post-transition than non-working parent controls $\hat{\gamma}_{31} = 0.04$, 95% CI [0.02, 0.06], $p < .001$; suggestive evidence in the nonparent sample: $\hat{\gamma}_{31} = 0.03$, 95% CI [0.01, 0.05], $p = .015$). We found that grandparents providing grandchild care increased more strongly in openness than matched parent controls (difference in *after* parameter: $\hat{\gamma}_{21} + \hat{\gamma}_{31} = 0.04$, 95% CI [0.01, 0.06], $p = .005$; suggestive evidence in the nonparent sample: $\hat{\gamma}_{21} + \hat{\gamma}_{31} = 0.03$, 95% CI [0.00, 0.05], $p = .025$). However, grandparents who provided grandchild care did not differ significantly from grandparents who did not (see Tables S50 & S51 and Figure S23). We found no evidence for

---

**Figure 4.** Change trajectories of agreeableness based on the basic models (Left column) and the models including the gender interaction (Right column). The error bars are 95% confidence intervals of the predicted values, which only account for the fixed-effects portion of the model. The vertical line indicates the approximate time of the transition to grandparenthood.
moderation of openness by ethnicity (see Tables S52 & S53 and Figure S24).

**Life satisfaction.** We found no consistent evidence that grandparents’ life satisfaction trajectories differed significantly from those of the controls in either the basic models (see Tables S54 & S55 and Figure S25) or the models including the gender interaction (see Tables S56 & S57 and Figure S25). There was also no evidence of a moderation of life satisfaction by performing paid work (see Tables S58 & S59 and Figure S26) or grandchild care (see Tables S60 & S61 and Figure S27).

Black/African American grandparents increased to a higher degree in life satisfaction after the transition to grandparenthood than Black/African American non-parent controls (difference in after parameter: \(\gamma_{41} + \gamma_{51} = 0.37, 95\% CI [0.14, 0.59], p = .001\); suggestive evidence in the parent sample: \(\gamma_{41} + \gamma_{51} = 0.28, 95\% CI [0.06, 0.50], p = .013\); see Tables S62 & S63 and Figure S28). In addition, there was suggestive evidence that Black/African American grandparents’ post-transition increases were more pronounced than those of White grandparents (difference in after parameter; parents: \(\gamma_{50} + \gamma_{51} = 0.28, 95\% CI [0.07, 0.49], p = .009\); non-parents: \(\gamma_{40} + \gamma_{51} = 0.29, 95\% CI [0.08, 0.49], p = .006\)). However, the model uncertainty regarding these effects was comparatively high.

**Interindividual differences in change**

First, we conducted model fit comparisons between the random intercept models reported previously and models where a random slope variance was estimated, separately for each change parameter because joint random effects modeling frequently led to model nonconvergence. These comparisons showed a substantial amount of interindividual differences.
differences in change for all random slopes in all models, as indicated by increases in model fit significant at \( p < .001 \).

Second, we estimated models with heterogeneous random slope variances of the grandparents and each control group to test whether interindividual differences in change were significantly larger in the grandparents. Contrary to hypothesis H2, for agreeableness, conscientiousness, extraversion, and neuroticism,
interindividual differences in intraintividual change were greater in the control group for all tested effects (see Tables S64, S65, S66, & S67). In the two HRS samples, assuming group heterogeneity in the random slope variances led to significant improvements in model fit in all model comparisons. In the two LISS samples, this was the case for around half the tests.

For openness, interindividual differences in change before the transition to grandparenthood were significantly greater in the HRS grandparents than the nonparent controls (random slope variances of the before parameter), likelihood ratio = 57.57, p < .001. This result could not be replicated in the other three samples. The other parameters of change either did not differ between groups in their random slope variances or had significantly larger random slope variances in the respective control group (see Table S68).

We found larger interindividual differences in grandparents’ changes in life satisfaction before the transition to grandparenthood compared to the nonparent controls in the HRS (random slope variances of the before parameter), likelihood ratio = 115.87, p < .001 (see Table S69). This was not corroborated in the other three analysis samples. Overall, most tests for heterogeneous random slope variances in life satisfaction indicated either non-significant differences or significantly larger random slope variances in the control sample.

### Table 4. Fixed effects of conscientiousness over the transition to grandparenthood moderated by grandchild care.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Parent controls</th>
<th>Nonparent controls</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\hat{\beta}$</td>
<td>95% CI</td>
</tr>
<tr>
<td>Intercept, $\gamma_{00}$</td>
<td>3.43</td>
<td>[3.39, 3.47]</td>
</tr>
<tr>
<td>Propensity score, $\gamma_{02}$</td>
<td>0.03</td>
<td>[-0.04, 0.10]</td>
</tr>
<tr>
<td>After-slope, $\gamma_{20}$</td>
<td>0.00</td>
<td>[-0.01, 0.01]</td>
</tr>
<tr>
<td>Grandparent, $\gamma_{01}$</td>
<td>0.01</td>
<td>[-0.05, 0.07]</td>
</tr>
<tr>
<td>Caring, $\gamma_{10}$</td>
<td>0.02</td>
<td>[-0.01, 0.06]</td>
</tr>
<tr>
<td>After-slope * Grandparent, $\gamma_{21}$</td>
<td>0.00</td>
<td>[-0.02, 0.02]</td>
</tr>
<tr>
<td>After-slope * caring, $\gamma_{30}$</td>
<td>-0.01</td>
<td>[-0.02, 0.00]</td>
</tr>
<tr>
<td>Grandparent * caring, $\gamma_{11}$</td>
<td>-0.06</td>
<td>[-0.14, 0.02]</td>
</tr>
<tr>
<td>After-slope * Grandparent * caring, $\gamma_{31}$</td>
<td>0.04</td>
<td>[0.01, 0.07]</td>
</tr>
</tbody>
</table>

Note. Two models were computed (only HRS): grandparents matched with parent controls and with nonparent controls. CI = confidence interval. caring = 1 indicates more than 100 hours of grandchild care since the last assessment.

### Figure 6. Change trajectories of conscientiousness based on the models of moderation by grandchild care (see Table 4). The error bars are 95% confidence intervals of the predicted values, which only account for the fixed effects portion of the model. The plots in the left column are the same as in Figure S11 (basic models) but restricted to the post-transition period for better comparability.
Findings: We conducted a preregistered, cross-study, and multi-comparison test of the social investment principle (Lodi-Smith & Roberts, 2007; Roberts & Wood, 2006) with grandparenthood as a candidate catalyst of personality change (Hutteman et al., 2014). We found more evidence of trait stability than of change. The direction of the few effects we found generally supported the social investment principle, that is, increases to agreeableness and conscientiousness and decreases to neuroticism—in contrast to development following parenthood (Asselmann & Specht, 2020b; van Scheppingen et al., 2016). However, even though small psychological effects may be meaningful and involve real-world consequences (Götz et al., 2021), the effects we found were not only small but also inconsistent across analysis samples.

Discussion
In an analysis of first-time grandparents compared to both parent and nonparent matched control respondents, we found pronounced stability in the Big Five and life satisfaction over the transition to grandparenthood. There were a few isolated effects in line with our hypotheses on mean-level increases in agreeableness and conscientiousness and decreases in neuroticism (H1a). However, they were very small in size, only present in grandfathers, and not consistent over the two analyzed panel studies (LISS and HRS) or the two matched control groups (parents and nonparents). We found no robust evidence that grandparents providing grandchild care experienced more pronounced positive personality development than those who did not (H1b). Evidence for moderation of mean-level trajectories by performing paid work was inconsistent. There was no evidence that grandmothers (or grandfathers) reached higher levels of life satisfaction following the transition to grandparenthood (H1c). Although interindividual differences in change were present for all change parameters, they were only greater in the grandparents than the controls in a small minority of model comparisons (H2). Finally, rank-order stability did not differ between grandparents and either control group, or it was lower in the control group—contrary to expectations (H3).
provided initial evidence that more frequent contact with grandchildren is associated with higher grandparental well-being (Arpino, Bordone, et al., 2018; Danielsbacka et al., 2019; Danielsbacka & Tanskanen, 2016; Dunifon et al., 2020). However, Danielsbacka et al. (2019) noted that this effect is due to between-person differences in grandparents, thus, limiting a causal interpretation of frequency of grandchild care as a mechanism of development in psychological characteristics like life satisfaction and personality.

**Life satisfaction**

Similar to the Big Five personality traits, we did not find convincing evidence that life satisfaction changed due to grandparenthood. A study of the effects of the transition on first-time grandparents’ life satisfaction that used fixed effects regressions also did not discover any positive within-person effects of the transition (Sheppard & Monden, 2019; see also Ates, 2019). Further, in line with this study, we did not find evidence that grandparents who provided grandchild care increased more strongly in life satisfaction than those who did not, and grandparents’ life satisfaction trajectories were also not moderated by employment status (Sheppard & Monden, 2019).

Overall, evidence has accumulated that there is an association between having grandchildren and higher life satisfaction on the between-person level—especially for (maternal) non-coresiding grandmothers who provide grandchild care (Danielsbacka et al., 2011, 2022; Danielsbacka & Tanskanen, 2016)—but no within-person effect of the transition. The main reason for this divergence is the presence of selection effects. Specifically, through propensity score matching we controlled for confounding (Luhmann et al., 2014; Thoemmes & Kim, 2011; VanderWeele et al., 2020), but its influence was present in previous studies. We carefully deliberated the inclusion of each covariate on the basis of its assumed causal relations to treatment assignment and the outcomes and made these underlying assumptions transparent within the preregistration.

In an exploratory analysis, Black/African American grandparents—usually lower in life satisfaction compared to White HRS respondents (e.g., Zhang et al., 2017)—increased in life satisfaction following the transition to grandparenthood bringing them up on par with White respondents. This is in line with cross-sectional data indicating no ethnic differences in life satisfaction between African American and White grandmothers (Goodman & Silverstein, 2006). Corroboration of this tentative finding in other samples should be awaited, though.

**Interindividual differences in change**

All parameters of change exhibited considerable interindividual differences. Similar to Denissen et al. (2019), who found model fit improvements with random slopes in most models (see also Doré & Bolger, 2018), respondents—both grandparents and matched controls—deviated to a considerable extent from mean-level change trajectories.
We expected larger interindividual differences in grandparents because life events differ in their impact on daily life and in the degree to which they are perceived as meaningful or emotionally significant (Doré & Bolger, 2018; Luhmann et al., 2020). Another reason for expecting heterogeneity in the individual trajectories were the considerable differences between grandparents in the amount of grandparental investment (e.g., Danielsbacka et al., 2022) and competing role demands (e.g., Arpino & Bellani, 2022) present in our samples. Our results, however, indicated that interindividual differences were larger in the controls than the grandparents for many models, or not significantly different between groups. Only in a small minority of tests were interindividual differences significantly larger in grandparents (concerning the linear slope in anticipation of grandparenthood for openness and life satisfaction).

Importantly, most previous studies do not compare interindividual differences in personality change between an event group and a comparison group (even if they use comparison groups for the main analyses; Denissen et al., 2019; Schwaba & Bleidorn, 2019; cf. Jackson & Beck, 2021). Interindividual differences in personality change are substantial up until around 70 years of age (Schwaba & Bleidorn, 2018). Regarding the substantive question of how the transition to grandparenthood affects interindividual differences in change, we propose that it is more informative to test grandparents’ variability in change against well-matched control groups than against no groups.

Recently, Jackson and Beck (2021) presented evidence that the experience of 16 commonly analyzed life events was mostly associated with decreases in interindividual variation in the Big Five compared to those not experiencing the respective event. They used a comparable approach to ours but in a SEM latent growth curve framework and without accounting for pre-existing group differences (i.e., without matching). Their results based on the German SOEP data suggested—contrary to their expectations—that most life events made people more similar to each other (Jackson & Beck, 2021). Thus, taken together with our results, it seems that the assumption that life events and transitions ostensibly produce increased heterogeneity between people needs to be scrutinized in future studies. It is possible that normative social demands of events such as grandparenthood increase homogeneity of personality development trajectories.

**Rank-order stability**

We expected lower rank-order stability over the transition to grandparenthood in grandparents compared to the matched controls based on the assumption that grandparents’ personality is reorganized through the experience of the event and the addition of the new social role. Conceptually, rank-order stability represents to which extent individual differences endure over time and it can be low even in the absence of mean-level changes if traits change nonsystematically. Empirically, though, we did not find evidence supporting our hypothesis (H3): Rank-order stability was highly similar in most comparisons of grandparents and controls, and it was not significantly lower in these comparisons. In a recent study of the effects of eight different life events on the development of the Big Five personality traits and life satisfaction (Denissen et al., 2019), comparably high rank-order stability was reported in the event groups. Only particularly adverse events such as widowhood and disability significantly lowered rank-order stability (Chopik, 2018; Denissen et al., 2019).

Regarding the Big Five’s general age trajectories of rank-order stability, support for inverted U-shape trajectories was recently strengthened in a study of two panel data sets (Seiffert et al., 2021). This study also explored that health deterioration accounted for parts of the decline of personality stability in old age. Therefore, it is possible that in later developmental phases (see also Hutteman et al., 2014) rank-order stability of personality is largely influenced by health status and less by normative life events. In the context of grandparenthood, this relates to research into health benefits (Chung & Park, 2018; Condon et al., 2018; Di Gessa et al., 2016a, 2016b; cf. Ates, 2017) and decreases to mortality risk associated with grandparenthood or grandchild care (Choi, 2020; Christiansen, 2014; Hilbrand et al., 2017; cf. Ellwardt et al., 2021). Grandparenthood might therefore have a time-lagged effect on personality stability through protective effects on health. However, with the currently available data, such a mediating effect cannot be reliably recovered (under realistic assumptions; Rohrer et al., 2022).

**Limitations and future directions**

A number of limitations need to be addressed: First, there remains some doubt whether we were able to follow truly socially invested grandparents over time. The moderator variable on grandchild care only reflects whether a respondent (or their spouse/partner) provides a minimal level of care. More detailed information regarding a grandparent’s relationship with their first and later grandchildren10 and the level of care a grandparent provides would be a valuable source of information on social investment, as would information on constraining factors such as length and cost of travel between grandparent and grandchild. One way to obtain comprehensive information on mechanisms of grandparental development would be a measurement burst design in a sample of grandparents with diverse social backgrounds (see Crawford et al., 2022; Springstein et al., 2022). This would allow differentiating contexts of social investment while also providing insight into daily-life social activities (e.g., Dunifon et al., 2020) and their medium- to long-term influence on personality development (Wrezus & Roberts, 2017). On a similar note, we did not examine grandparents’ subjective perception of the transition to grandparenthood in terms of the emotional significance, meaningfulness, and impact on daily lives, which might be responsible for differential individual change trajectories (Haehner et al., 2022; Kritzler et al., 2022; Luhmann et al., 2020). Grandparents’ perception of potential role conflicts (Goode, 1960), and whether they perceive caregiving as a burden or obligation (Xu et al., 2017), could also uncover mechanisms of personality development.

Second, a causal interpretation of our results rests on a number of assumptions that are not directly testable with the data (Li, 2013; Stuart, 2010): We assumed that we picked the right sets of covariates, that our model to estimate the propensity score
was correctly specified, and that there was no substantial remaining bias due to unmeasured confounding. Importantly, we selected covariates following state-of-the-art recommendations and substantiated each covariate’s selection explicitly within our preregistration. Regarding the propensity score estimation, we computed grandparents’ propensity scores at a specific time point at least 2 years before the transition to grandparenthood, which had the advantages that (1) the covariates were uncontaminated by anticipation of the transition, and (2) the matched controls had a clear counterfactual timeline of transition (for similar approaches, see Balbo & Arpino, 2016; Krämer & Rodgers, 2020; van Scheppingen & Leopold, 2020). It also has to be emphasized that the timing of measurements might have missed more short-term effects of grandparenthood playing out over months instead of years.

Third, our results only pertain to the countries for which our data are representative on a population level: the Netherlands and the United States. Personality development has been examined cross-culturally (e.g., Bleidorn et al., 2013; Chopik & Kitayama, 2018): On the one hand, these studies showed universal average patterns of positive personality development over the life span. On the other hand, they emphasized cultural differences regarding norms and values and the temporal onset of social roles (Arshad & Chung, 2022). For grandparenthood, there are demographic differences between countries (Leopold & Skopek, 2015), as well as differences in public child care systems that may demand different levels of grandparental involvement (Bordone et al., 2017; Hank & Buber, 2009). In the Netherlands, people become grandparents 6 years later on average than in the United States (Leopold & Skopek, 2015). Furthermore, although both countries have largely market-based systems for early child care, parents in the Netherlands on average have access to more extensive childcare services through (capped) governmental benefits (OECD, 2020). Despite these differences, our results from the Dutch and US samples did not indicate systematic discrepancies.

Conclusion

Do personality traits change over the transition to grandparenthood? In two nationally representative panel studies in a preregistered propensity score matching design, Big Five personality traits and life satisfaction remained predominantly stable in first-time grandparents over this transition compared to matched parents and nonparents. We found slight post-transition increases in grandparents’ agreeableness and conscientiousness in line with the social investment principle. However, these effects were minuscule and inconsistent across analysis samples. In addition, our analyses revealed (1) a lack of consistent moderators of personality development, (2) interindividual differences in change that were mostly smaller in grandparents than in matched respondents, and (3) comparable rank-order stability in grandparents and matched respondents. Thus, we conclude that the transition to grandparenthood did not act as a straightforwardly important developmental task driving personality development (as previously proposed, see Hutteman et al., 2014). With more detailed assessment of the grandparent role, future research can investigate whether personality development occurs in grandparents with specific degrees of role investment.

Acknowledgments

We thank Joe Rodgers, Jaap Denissen, Oliver Huxhold, and Julia Rohrer for helpful comments on earlier versions of this paper.

Authors’ contribution

The authors made the following contributions. Michael D. Krämer: Conceptualization, Data Curation, Formal Analysis, Methodology, Visualization, Writing—Original Draft Preparation, Writing—Review and Editing; Manon A. van Scheppingen: Methodology, Writing—Review and Editing; William J. Chopik: Methodology, Writing—Review and Editing; David Richter: Supervision, Methodology, Writing—Review and Editing.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Data accessibility statement

This article earned Open Data, Preregistered and Open Materials badges through Study materials, the preregistration, and scripts for data wrangling, analyses, and to reproduce this manuscript can be found on the OSF (https://osf.io/75a4r/) and GitHub (https://github.com/mdkraemer/gp-personality). We also provide instructions to aid reproducing the manuscript. LISS and HRS data are available online to registered users.

ORCID iD

Michael D Krämer https://orcid.org/0000-0002-9883-5676

Supplemental Material

Supplemental material for this article is available online.

Notes

1. Fixed effects regression models rely exclusively on within-person variance (see Brüderl & Ludwig, 2015; McNeish & Kelley, 2019).

2. Dichotomization of a continuous construct (hours of care) is not ideal for moderation analysis (MacCallum et al., 2002). However, there were too many missing values in the variable assessing hours of care continuously (variables *E063).

3. We also excluded N = 30 HRS grandparents in a previous step who reported unrealistically high numbers of grandchildren (>10) in their first assessment following the transition to grandparenthood.

4. In these logistic regressions, we included all covariates listed above as predictors except for female, which was later used for exact matching, and health-related covariates in LISS wave 2014, which were not assessed in that wave.

5. In the LISS, 282 grandparent observations were matched with 1128 control observations; these control observations corresponded to 561 unique person-year observations stemming from 281 unique respondents for the parent control group, and to 523 unique person-year observations stemming from 194
unique respondents for the nonparent control group. In the HRS, 847 grandparent observations were matched with 3388 control observations; these control observations corresponded to 1363 unique person-year observations stemming from 978 unique respondents for the parent control group, and to 1039 unique person-year observations stemming from 712 unique respondents for the nonparent control group.

6. We also provide instructions to aid reproducing the manuscript.

7. As a robustness check, we re-estimated the mean-level trajectories after further restricting the time frame by excluding time points earlier than 2 years before the transition (i.e., before the latest time of matching). This served the purpose of assessing whether including time points from before matching (as preregistered) would distort the trajectories in any way. However, results were highly similar (see gp_restricted_models.pdf on osf.io/75a4r).

8. In addition to the preregistered retest interval, we computed a maximally large interval between the first available assessment before and the last assessment after the transition. Here, 3 out of 24 comparisons indicated that rank-order stability was lower in the grandparents. There was only one significant difference supporting our hypothesis: HRS grandparents’ rank-order stability in openness was lower than that of the nonparents, \( p < .001 \) (see Table S70). Another analysis also failed to provide convincing evidence that grandparents’ rank-order stability was lower: We excluded duplicate control respondents resulting from matching with replacement who might bias results towards greater stability in the controls. Descriptively, 10 out of 24 comparisons showed lower rank-order stability in the grandparents (see Table S71). However, group differences were small and non-significant.

9. In the HRS, a higher proportion of first-time grandmothers (M = 0.45, SD = 0.50) than grandfathers (M = 0.41, SD = 0.49) reported that they provided at least 100 hours of grandchild care since the last assessment.

10. It is also possible that effects of grandparental role investment accumulate with successive grandchildren (as shown for parental sleep deficits; Richter et al., 2019).

References


Asselmann, E., & Specht, J. (2020a). Taking the ups and downs at the rollercoaster of love: Associations between major life events in the domain of romantic relationships and the Big Five personality traits. Developmental Psychology, 56(9), 1803–1816. https://doi.org/10.1037/dev0001047


Austin, P. C. (2017). Double propensity-score adjustment: A solution to design bias or bias due to incomplete matching.
Krämer et al. 581


Margolis, R., & Wright, L. (2017). Healthy grandparenthood: How long is it, and how has it changed? Demography, 54(6), 2073–2099. https://doi.org/10.1007/s13524-017-0620-0


