UNIVERSITY of CALIFORNIA PRESS | Collabra: Psychology

Rohrer, J., Seifert, I. S., Arslan, R. C., Sun, J., & Schmukle, S. C. (2024). The Effects of Satisfaction With Different Domains of Life on General Life Satisfaction Vary Between Individuals (but We Cannot Tell You Why). *Collabra: Psychology*, *10*(1). https://doi.org/10.1525/collabra.121238

Personality Psychology

The Effects of Satisfaction With Different Domains of Life on General Life Satisfaction Vary Between Individuals (but We Cannot Tell You Why)

Julia Rohrer[1a], Ingo S. Seifert[1], Ruben C. Arslan[1], Jessie Sun[2], Stefan C. Schmukle[1]

[1] Wilhelm Wundt Institute for Psychology, Leipzig University, Leipzig, Germany, [2] Department of Psychological & Brain Sciences, Washington University in St. Louis, , MO, US

Keywords: life satisfaction, domain satisfaction, importance weighting, diary study

https://doi.org/10.1525/collabra.121238

Collabra: Psychology

Vol. 10, Issue 1, 2024

People care about different domains of life (e.g., their health, social life, work) to varying degrees. It thus seems plausible that how satisfied they are with those domains matters for their general life satisfaction to varying degrees. This idea has been investigated in the importance-weighting literature with at best mixed results, but variations of it can be found across different fields of psychology and include claims that values, personality, and age moderate the extent to which different life domains affect life satisfaction. In this study, we investigated the effects of satisfaction with 14 different life domains on general life satisfaction in a study of 439 individuals who provided up to 15 diary entries, resulting in a total of 6,071 observations. All domains had positive effects on average, with the largest effects for satisfaction with leisure time usage (b = 0.19, b_{std} = 0.25 relative to the within-person variability) and relationship satisfaction (b = 0.16, b_{std} = 0.17). Beyond these averages, there was robust interindividual variability; the standard deviation of the individual-level effects was of a similar magnitude as the average effect (and sometimes even larger). But when exploring correlations between these individual-level effects with third variables (e.g., self-reported importance of the respective domain, gender and age, Big Five personality traits), no convincing overall patterns arose. This may at least in part result from the high uncertainty with which individual-level effects were estimated, with reliabilities of ~.30, and the resulting low statistical power.

What makes for a satisfying life? This question has spawned a large number of empirical studies and theoretical debates across the social sciences (e.g., Diener et al., 2018). From everyday experience, but also the perspective of interindividual difference research, it seems obvious that at least part of the answer has to be “it depends on the person.” For some, their family life may be of tremendous importance for their well-being; for others, friends may matter more. For some, work may be a major source of life satisfaction; for others, leisure may receive a much higher weight. A better understanding of such interindividual differences would not only help us paint a more nuanced picture of life satisfaction; it would also be of interest from various theoretical perspectives. For example, individual difference researchers may be interested in how personality affects what matters for well-being (e.g., Gerson et al., 2016; Heller et al., 2006; Kinnunen et al., 2003; Węziak-Białowolska et al., 2019); developmental researchers may want to trace how the influence of different factors changes over the life course (e.g., Böger & Huxhold, 2018; Huxhold et al., 2014; Potter et al., 2021; Schöllgen et al., 2016).

In this article, we investigate such interindividual differences in the determinants of well-being following a mostly exploratory approach. More specifically, we aim to get a better understanding of interindividual differences in the effects of domain satisfaction (e.g., satisfaction with health, social life, work, and leisure) on general life satisfaction. Our data comes from a longitudinal diary study, which allows us to, in the first step, estimate individual-level effects and quantify their heterogeneity. In contrast to previous studies, we explicitly aim to identify causal effects and spell out the necessary assumptions for their identification. If interindividual differences in the effects of domain satisfaction exist, we would want to be able to explain them, and so in the second step, we correlate the estimated individual-level effects with various other variables. Previous studies have often focused on a few or even just a single life domain (e.g., social life), and how its effects on well-be-

a Corresponding author: julia.rohrer@uni-leipzig.de

Table 1. Overview of our Research Questions

No	Question	Motivation
1	How much interindividual variability is there in the individual-level effects of satisfaction with different life domains (see Table 2) on general life satisfaction?	Establishing the presence of meaningful interindividual variability before trying to explain it.
2	How reliably can we estimate these effects?	Establishing that our subsequent analyses are meaningful; providing recommendations for future studies.
3	Do these effects correlate with individuals' average domain satisfaction, or their variability in domain satisfaction?	Ruling out that interindividual variability in the effects of domain satisfaction arises "mechanistically" from features of the individual-level domain satisfaction distribution.
4	Do these effects correlate with age and gender?	Exploring whether there could be differences in the effects depending on demographic characteristics.
5	Do these effects correlate with explicit assessments of the role of the domain?	Exploring whether our data provide any evidence for the idea of importance weighting.
6	Do these effects correlate with the Big Five personality traits?	Exploring whether the individual effects could be reflective of personality traits.

ing may covary with another individual variable (e.g., age; Huxhold et al., 2014). Here, we try to provide a broader picture by including a total of 14 domain satisfaction ratings, and by exploring associations with a range of other variables including gender, age, the Big Five, and also participants' assessment of the importance of the domains (see Table 1 for an overview of our research questions [RQs], presented in the order in which results are discussed).

Research on Interindividual Differences in Effects on Well-Being

The idea that individuals vary in how strongly their well-being is affected by their everyday lives can be found across multiple literatures in psychology. These invoke distinct theoretical frameworks to motivate their hypotheses and use different terms to describe such variation; however, the underlying structure of the claims is the same: the effect of X (something happening in individual's lives) on Y (some well-being outcome) varies (RQ 1 in Table 2), and such variation is often explained by some third variable M, a so-called moderator (RQs 4-6).

Most directly relevant to our work is the literature on importance weighting in subjective well-being research (RQ 4, see, e.g., Campbell et al., 1976; Hsieh & Li, 2020; Rohrer & Schmukle, 2018). Importance weighting refers to the idea that general life satisfaction is a weighted aggregate of satisfaction with various life domains (e.g., family, work, health), in which the contribution (i.e., weight) of any single domain depends on how important it is to the individual. This importance is captured by ratings provided by the individual (e.g., "How important is your family life to you?" answered on a rating scale). In other words, the idea is that the effects of domain satisfaction on general life satisfaction vary, and that variation is explained by the reported importance of the respective life domain. The idea goes back as far as seminal work on life satisfaction conducted by Campbell et al. (1976), but despite a wealth of empirical studies on the topic, there is little consensus in the literature (literature reviewed in e.g., Rohrer & Schmukle, 2018). The repeated failures to find evidence for importance weighting may of course indicate that it is just substantively wrong, but they may also result from a wealth of methodological issues. For example, researchers routinely use single-item importance ratings (e.g., "How important is your family life to you?"). These often show little variability (i.e., almost everybody will say that their family life is very important), and they may also lack reliability or validity (see, e.g. Hsieh & Li, 2020; Rohrer & Schmukle, 2018, for discussions of these and other issues). Furthermore, findings may also hinge on which domains are included in the analysis, a concern to which we will return later when discussing causal identification assumptions.

Problems with the Reliance on Cross-Sectional Observational Data

One methodological concern that has been neglected in this particular literature is the fact that it relies on cross-sectional, observational data. The question of importance weighting is a causal one: It is, at the very least, assumed that domain satisfaction causally affects general satisfaction (if domain satisfaction is changed, general satisfaction changes). Additionally, importance weighting may be interpreted to imply that the magnitude of these effects is causally affected by domain importance (if domain importance is changed, the effect of domain satisfaction changes, i.e., a causal interaction); we will leave aside this aspect for now and focus on the identification of the effects of domain satisfaction, and interindividual differences in such effects.

While cross-sectional observational data *can* be informative for causal claims, the necessary assumptions regarding the absence of unobserved confounding are often rather heroic (see e.g., Rohrer, 2018). In the importance weighting literature, potential confounding is usually completely ignored. Consider the question of whether the reported importance of family moderates the effect of family life satisfaction on general satisfaction. Possible confounders may include variables such as gender, age, the presence of children, how close other family members are living, personality, and religiosity. These potentially introduce spurious associations between family life satisfaction and general satisfaction. But what does this mean for the question of moderation by importance?

In principle, even in the presence of confounding, one can imagine a scenario in which moderation by family importance still successfully recovers differences in the causal effects of family life satisfaction on general satisfaction. In this scenario, all confounding is linear and additive across different values of importance: if we split the sample according to the importance variable, then in every subsample the confounding between family satisfaction and general satisfaction will be the same. Thus, if we compare the (confounded) effect estimates between different levels of importance, the confounding "subtracts out," leaving only differences in the actual causal effects. But the assumption that any confounding is linear and additive is extremely restrictive and usually not justified. Usually, it is assumed that confounding may vary between levels of the moderator, which necessitates that controls are interacted with the independent variable of interest (Rohrer & Arslan, 2021; Simonsohn, 2019; Yzerbyt et al., 2004). If no controls are included to begin with, as is the case in the importance weighting literature, such varying confounding necessarily remains unmodelled.

Thus, if the effects that are supposed to be moderated (e.g., the effects of domain satisfaction on general satisfaction) are plausibly confounded, we have to assume that the moderation results do not reflect differences in the effect of interest, but rather an uninterpretable blend of differences in causal and non-causal associations. Unfortunately, the importance weighting literature is thus likely rather uninformative when it comes to the question of whether the effects of various aspects of life on well-being vary between individuals.

This problem is by no means unique to the importance weighting literature. For example, there is a wealth of cross-sectional observational studies investigating whether personality moderates the effects of various possible causes of well-being (RQ 6). Does emotional stability moderate the effects of work-family conflict on well-being (Kinnunen et al., 2003)? Does personality moderate the effects of social comparison on subjective well-being (Gerson et al., 2016)? Does it moderate the effects of participation in cultural events (Węziak-Białowolska et al., 2019)? These studies often report no plausible causal identification strategy, thus leaving it unclear whether much can be learned about interindividual differences in the causes of well-being.[1]

Benefits of Longitudinal Studies

At this point, longitudinal data provide a productive way forward. While observational longitudinal data do not automatically enable causal claims, they can help relax some of the necessary assumptions. In particular, if analyzed properly, longitudinal models can control for the effects of time-invariant confounders—confounders whose value is constant throughout the study—regardless of whether said confounders have been measured or not. They achieve so by removing between-person differences, thus only leaving within-person associations that cannot be attributed to time-invariant confounders (Rohrer & Murayama, 2023). Thus, depending on the time covered by the study, we no longer need to worry about the confounding influences of, for example, gender, age, household composition, socioeconomic status, stable personality traits, values, and preferences. We thus can identify the effects of domain satisfaction on general satisfaction under the somewhat weaker assumption of no unobserved time-varying confounding, which in turn makes it more plausible that we can identify interindividual differences in such effects.

We are not aware of any studies analyzing longitudinal data to address the question of importance weighting in the narrower sense; however, there are individual longitudinal studies in the personality literature relevant to the question of whether there are interindividual differences in the causes of life satisfaction (RQ 6). For example, Heller et al. (2006) hypothesized that more neurotic individuals would be more sensitive to changes in job satisfaction and changes in marital satisfaction but failed to find evidence for said prediction in a longitudinal study.

Values as Moderators

One line of literature in the field of personality research that has relied on longitudinal data investigates the idea of values as moderators. People vary considerably in what they consider part of a good life (Pfund et al., 2024; Willroth et al., 2024). Oishi et al. (1999) suggested a model in which individuals' values (e.g., achievement, conformity, benevolence) influence their sources of subjective well-being. Thus, the idea is once again that the effects of various factors on well-being vary; this time the variation is explained by values. In a diary study of college students, Oishi et al. predicted general life satisfaction and found statistically significant interaction terms between grade satisfaction and achievement values, between family satisfaction and conformity values, and between social life satisfaction and benevolence values, which they took as evidence for the suggested values-as-moderators model. The authors further discussed the conceptually intriguing idea that developmental stages may affect these values and thus shape qualitative aspects of subjective well-being through a particular mechanism (RQ 4).

This developmental angle was picked up by Cheung and Lucas (2015), who investigated how the association be-

1 They may, at best, help us predict what predicts well-being in certain people. But such prediction will always be conditional on any other variables included in the analysis and on the population being investigated, and it may be of little practical use. First, the added predictive utility of including moderation will often be small. Second, we rarely encounter situations in which we do have information on determinants of well-being and of the involved moderators, but we do not have information on well-being, so that it makes sense to use the former to infer the latter.

tween income and life satisfaction changes across the life course in a longitudinal design. Using panel data from Germany, Great Britain, and Switzerland, they found that the within-person association between income and life satisfaction was strongest in midlife. They then investigated whether the moderation by age could partially be explained (i.e., mediated) by family values, and found statistical evidence compatible with such a mediated moderation in line with the values-as-moderators model.

Age as Moderator

Age-related differences in the effects of various factors on well-being (RQ 4) are also the focus of various studies in aging research. Böger and Huxhold (2018) investigated (reciprocal) relationships between loneliness and negative affect. Their longitudinal data were compatible with the idea that the effect of loneliness on negative affect was smaller among older participants. In a similar vein, Huxhold et al. (2014) found, among other things, that only among middle-aged individuals (and not among older individuals), activities with family members had positive effects on life satisfaction.

Under the label of "health sensitivity," aging researchers investigate whether changes in health are more or less important for well-being among older individuals. For example, Schöllgen et al. (2016) investigated within-person associations between functional limitations (a proxy for physical health) and depressive affect (an indicator of low well-being) and found that they decreased with age. Potter et al. (2021) found that older adults had a smaller negative within-person association of physical symptoms and positive affect compared to younger adults; but such a decreased health sensitivity was not found for negative affect.

While longitudinal data are more promising concerning causal inference, unfortunately, the published studies on heterogeneity in the effects of life domains on well-being usually do not address causality explicitly. Instead, the focus is on the statistical model and causal inference happens "between the lines" (as is customary in psychology, Grosz et al., 2020). Thus, it remains unclear whether/which causal effects are targeted, and under which assumptions the studies can successfully identify heterogeneity in causal effects.

The Present Study

In sum, there are many cross-sectional observational studies investigating and trying to explain heterogeneity in the effects of various factors on well-being; however, the extent to which such studies can inform us about heterogeneity in causal effects is questionable. In contrast, longitudinal studies are more promising from a causal inference perspective, at least in principle. Here, individual studies exist that zoom in on specific factors potentially contributing to well-being and specific third variables that may correlate with and potentially explain such heterogeneity. However, given that these studies usually do not explicitly address questions of causality, it remains unclear what can be learned about heterogeneity in causal effects.

With our study, we want to contribute to a more comprehensive and systematic understanding of interindividual differences in the determinants of well-being. For this purpose, we aim to identify the causal effects of satisfaction with various life domains on general life satisfaction, quantify interindividual differences in the magnitude of these effects (RQs 1-2 in Table 2), and in a subsequent step probe potential explanations for such differences (RQs 3-6). Our focus on domain satisfaction (rather than more detailed assessments of what is going on within different life domains) stems from the fact that our project took the importance weighting literature as a starting point. It seems plausible that domain satisfaction ratings subsume the effects of many smaller factors within the respective domains and thus, using such ratings allows us to cover a broad range of aspects of life in a parallel manner.

In contrast to existing longitudinal studies on the question of heterogeneous effects on well-being; we (1) explicitly aim for the causal identification of the effects of domain satisfaction on general satisfaction and discuss the necessary assumptions and (2) add a preliminary step in which we aim to quantify interindividual differences in the effects and their reliability (RQs 1-2), rather than immediately skipping to potential explanations for why people vary.[2] Presenting heterogeneity without providing any sort of potential explanation is, of course, intellectually not quite as satisfying; we thus additionally explore whether the effects of domain satisfaction on general satisfaction correlate with various other variables: average and variability of domain satisfaction ratings (which may point to "artifactual" effect interindividual differences resulting from measurement issues; RQ 3), basic demographics (age and gender; RQ 4), respondents' ratings of the domains' importance and influence (which may provide evidence for importance weighting; RQ 5), and personality (the Big Five; RQ 6).

Method

Design of the Diary Study

General Structure

Individuals participated in a longitudinal online study administered with the help of formr (Arslan et al., 2020); Figure 1 provides an overview of the study design. Central research questions and exclusion criteria were pre-regis-

2 Our diary study was originally planned (and preregistered) as a (confirmatory) test of importance weighting. However, while working on the analyses, we noticed conceptual obstacles regarding the causal identification of the effects of interest. Thus, we shifted the focus from a narrow, potentially premature hypothesis test (Scheel et al., 2020) to a more descriptive and exploratory approach.

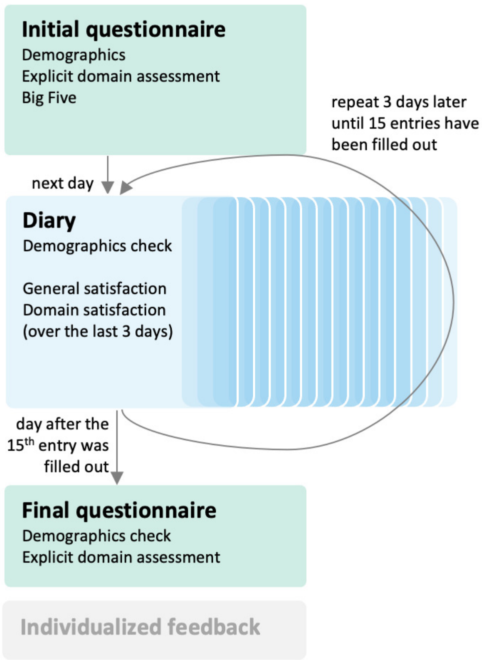

Figure 1. Overview of the Data Collection for the Diary Study

tered at an early stage of data collection, see https://osf.io/jpxmn. At the time of the preregistration, however, we were not sure how to best analyze the data, and explicitly stated so in the preregistration. Thus, the specific analyses (which we justify in much more detail below) should be evaluated as if the data analyses had not been preregistered. Deviations from the preregistration are noted throughout the manuscript and are also summarized in a preregistration deviations table (following recommendations by Willroth & Atherton, 2024) provided in the online supplement on the OSF (https://osf.io/x8j7r/).

The study started with an initial questionnaire collecting basic demographic information, explicit assessments of various life domains, as well as measures of the Big Five (see details below). The day after participants had filled out the initial questionnaire, they received an invitation to fill out the first diary entry at 15:00 (3:00 p.m.). In the diary, we collected assessments of their current general life satisfaction, as well as of their current satisfaction with various life domains. Instructions explicitly referred to the last three days. We chose this long retrospective window because we expected that in some domains, there may not be meaningful daily changes. For example, satisfaction with one's tasks at work may not change on the weekend assuming one does not work on those days; exercise satisfaction may not show meaningful variability if one did not exercise (and did not plan to do so) on a given day. We considered a three-day window a reasonable compromise: long enough so that chances are high that something meaningful happened in most domains (e.g., each interval contained at least one day of the workweek, improving chances that one's primary activity was relevant); short enough not to strain participants' memory.

Participants were also able to indicate relevant changes in their baseline "demographics" information (in case they, e.g., started or ended a romantic relationship; the diary questionnaire was then adapted accordingly). Participants were able to fill out the diary within a 12-hour time window, until 3:00 (3:00 a.m.) the next day. Regardless of whether they filled out the diary questionnaire, they received the e-mail invitation for the next diary questionnaire three days after the previous invitation had been sent.

This cycle was repeated until (a) 15 diary questionnaires had been filled out or (b) 80 days passed, at which point we assumed a participant had lost interest in the study. Those participants who had completed 15 diary entries were invited to fill out the final questionnaire the day after they finished the last diary entry; this questionnaire again included a demographics check and assessments of the importance of various life domains. Afterwards, participants could opt in to receive automatically generated individualized feedback containing their satisfaction over the course of the study, visual displays of their associations between domain satisfaction and general life satisfaction, as well as feedback on their values on the Big Five personality traits.

Due to the design of the study, time intervals between the diary entries provided by the participants could vary. Considering the final sample underlying our main analysis, the average interval was M = 4.3 days, SD = 3.12 days. The majority of entries (74%) were provided three days after the previous entry (i.e., as intended according to the design) but a substantial number of entries were provided after one missed time window (i.e., 6 days after the previous entry; 17%), after two missed time windows (i.e., 9 days after the previous entry; 5%), or after three missed time windows (i.e., 12 days after the previous entry, 2%). On average, participants included in the final sample spent 55.1 days in the diary part of the study (SD = 12.2 days, *Median* = 54 days).

The study included other measures which were not analyzed for the purpose of this manuscript. A list including all measures can be found in the public preregistration of this study at https://osf.io/jpxmn; more details can be found in the study materials at https://osf.io/x8j7r/.

Recruitment

We recruited German-speaking participants through multiple channels. We advertised the study to students at Leipzig University and the University of Ulm; posted physical advertisements in Leipzig, Ulm, and Coburg; and invited the members of various Facebook groups. Furthermore, we distributed the study through the online platform PsyWeb (https://psyweb.uni-muenster.de), whose users can be expected to be somewhat more diverse with respect to age and level of educational attainment than student samples. Given these channels of distribution, and given that the primary incentive to participate in the study was the individualized feedback provided at the very end, the resulting sample is most likely biased towards individuals with a certain interest in psychology and their own well-being.

Inclusion Criteria

For our key analyses, as preregistered, we included participants who had completed the initial questionnaire and at least five diary questionnaires. A cut-off of five diary entries may provide a reasonable balance: including participants with fewer entries may result in the inclusion of some participants who were not fully interested in the study, or for whom intraindividual associations are unreliable to the point of being uninformative; including only participants with more entries or even complete data diminishes the sample size and may introduce further selection bias.

Deviations from Preregistration. We did not conduct additional robustness checks included in our preregistration in which we planned to apply different inclusion criteria (1a: all diary entries plus final questionnaire, 1b: all diary entries, 1c: at least one diary entry). In hindsight, these criteria appear suboptimal (1a, 1b: removal of 106–108 participants of which 56 had actually filled out ten or more diary entries; 1c: inclusion of 68 participants for whom the individual-level effects can only be estimated extremely unreliably), even more so because our research question shifted from an explicit test of importance weighting (which may result in a single conclusion whose robustness can then be questioned) to the identification of interindividual differences in effects and an exploration of their correlates (which results in more descriptive and more numerous findings to report).

We had furthermore pre-registered specific exclusion criteria for participants who indicated signs of careless responding (specifically, excessive repetitions of the same response). We decided against applying these exclusion criteria. For most of the included measures, at best a handful of participants met the pre-registered exclusion cut-off (e.g., n = 5 had given the same importance rating to more than 85% of life domains, n = 0 met the repetition cut-off for the Big Five measure). For measures that met the repetition cut-off more frequently, it is plausible that the response pattern does not reflect careless responding but rather other factors (e.g., genuine uncertainty in the ratings of how much domains affect life satisfaction leading to a default response; a ceiling effect in the general life satisfaction assessment in the diary questionnaire).

Final Sample

The final sample consisted of 439 individuals (65% women, 34% men, and 1% non-binary or not reported) who filled out a total of 6,071 diary questionnaires (M = 13.8 questionnaires per participant, SD = 2.6). The average age of participants was 40.1 years (SD = 16.6 years). The vast majority (84%) of participants indicated that they were involved in some regular primary activity; including work (full-time: n = 160, part-time: n = 57), college (n = 134), school (n = 2), apprenticeship (n = 2) or other (n = 13; e.g., volunteering, self-employment, parental leave). More than half of the participants (56%) were in a romantic relationship.

Some analyses (e.g., retest-stability of domain assessments) necessarily rely on the subsample who filled out the final questionnaire (and thus also all 15 diary questionnaires). This subsample consisted of 361 individuals (65% women, 34% men, 1% non-binary or not further specified; M_{Age} = 39.3 years, SD_{Age} = 17.4 years; 82% in regular primary activity; 55% in some sort of romantic relationship).

Measures

Demographics

Participants were asked to report their age in years and their gender (response option: woman, man, and "other" which was combined with an optional free-text field). Furthermore, there were three questions to decide which life domains applied to participants. First, participants could indicate that they were involved in some sort of romantic relationship; those who replied "yes" were able to further tick features that applied to the relationship (e.g., casual, committed, married, cohabitating, long distance). Second, participants were asked whether they were involved in any regular activity (such as school, university, apprenticeship, or a profession). Those who confirmed were then asked to specify their primary activity (German: Haupttätigkeit) using a list of response options (school, university, apprenticeship, work full-time, work part-time) with the option to specify something else in a free-text field. All following questions referring to the primary activity were adapted accordingly. For example, participants who reported to work were asked whether they were satisfied with work; participants who specified their primary activity in the free-text field were asked whether they were satisfied with "their primary activity (<free-text label provided by the respondent>)." Third, participants were asked whether they physically exercised regularly, at least once per week. Because the answers to these three questions could change over time, every subsequent questionnaire after the initial questionnaire (i.e., all diaries and the final questionnaire) started with a brief demographics check in which participants could either confirm or update their last answer.

Included Life Domains

To decide which life domains to include in our study, we relied on previous studies, but also on our own criteria. For example, in previous studies, there was often very little variability in importance ratings—almost everybody considers their family and their health very important (Rohrer & Schmukle, 2018). This lack of variability makes it hard to detect evidence for importance weighting. Thus, we intentionally tried to also include some domains of satisfaction that are not deemed important by everyone (namely the way one looks and physical exercise). Table 2 lists all domains and the number of participants from the final sample to whom the domain applied.

We included seven life domains that we assumed to apply to everyone: Health, social contacts in everyday life, family, friends, the way one looks, leisure time available, and how said leisure time was put to use. Additional six life domains applied only to subgroups of participants: romantic relationship, physical exercise, and primary activity,

Table 2. Overview of the Included Life Domains and the Corresponding Satisfaction Items

Life domain label	In the last three days, how satisfied were you with…	Applies to	*N*	*M*	*SD*	$SD_{Between}$	SD_{Within}	*ICC*
Health	…your health?	All	439	4.86	1.54	1.12	1.05	.49
Social life	…your everyday social contacts, all in all?	All	439	5.16	1.25	0.94	0.83	.52
Family	…the relationship with your family?	All	439	5.24	1.32	1.07	0.78	.62
Friends	…the relationship with your friends?	All	439	5.09	1.31	1.05	0.78	.62
Looks	…your looks?	All	439	4.66	1.35	1.12	0.74	.67
Leisure time	…the amount of leisure time you had?	All	439	5.07	1.57	1.12	1.10	.47
Leisure use	…how you used your leisure time?	All	439	4.95	1.53	1.05	1.11	.43
Relationship	…your romantic relationship?	In romantic relationship	248	5.24	1.44	1.12	0.90	.58
Exercise	…your sporting activities?	Exercises regularly	276	4.72	1.51	0.73	1.16	.37
Primary activity								
General	…your primary activity, all in all?	Has primary activity	369	4.83	1.39	1.12	0.82	.62
Tasks	…contents and task at your primary activity?	Has primary activity	369	4.78	1.39	1.12	0.84	.61
Performance	…your performance at your primary activity?	Has primary activity	369	4.77	1.45	1.13	0.92	.56
Social contacts	…social contacts at your primary activity?	Has primary activity	369	4.96	1.39	1.10	0.86	.59
Individual domain	…your own domain (<custom label>)?	Reported own domain	162	4.33	1.77	1.50	0.94	.70
General	*Modified Satisfaction with Life Scale*	All	439	4.89	1.28	0.98	0.82	.55

Note. $SD_{Between}$ refers to the standard deviation of the person-specific means of domain satisfaction; SD_{Within} refers to the standard deviation of domain satisfaction after person-specific means were subtracted. Squaring these two numbers and adding them up results in the squared (total) SD. SDs were calculated across all observations. ICC refers to the intraclass correlation coefficient, estimated based on the variance components from a one-way ANOVA.
All questions relating to the primary activity were adapted to the primary activity reported by the respondent, e.g., "…how satisfied were you with your work, all in all?", "…how satisfied were you with your performance at school?"

which we split into four questions (primary activity in general, contents and tasks, performance, social contacts). Finally, participants had the possibility to add their own domain in the initial questionnaire. They were presented with an overview of the domains applying to them and were asked if anything important was missing; those who replied "yes" were then able to add another domain with a custom label in a free-text field. This domain was included in all subsequent assessments alongside all other domains using the custom label. About a third of participants made use of this option reporting a diverse range of domains, including, for example, finances, hobbies, religion, pets, and side jobs.

General and Domain Satisfaction (Diary Questionnaire)

To measure general satisfaction with life, we included an adapted version of the Satisfaction with Life Scale (Diener et al., 1985). All five items were rephrased to refer to the last three days (e.g., "Over the course of the last three days, I was satisfied with my life") and participants answered on a seven-point scale ranging from *strongly disagree* to *strongly agree*. Across all diary entries of individuals included in the final sample, the internal consistency of the general satisfaction scale was high (α = .93); it remained high (α = .89) when only considering within-person deviations by subtracting person-specific means from each item. We averaged the five items to arrive at a measure of general satisfaction.

Participants also reported their satisfaction with each of the life domains that applied to them. The instructions explicitly referred to the last three days (see Table 2) and participants answered on a seven-point scale ranging from *very dissatisfied* to *very satisfied*.

The online supplement on the OSF (https://osf.io/x8j7r/) provides between- and within-person correlations between general satisfaction with life and all domain satisfaction items using the decomposition procedure implemented in the *psych* package (Revelle, 2017). Between-person correlations were consistently positive—people who tended to report high satisfaction with one aspect also tended to report high satisfaction with other aspects—and ranged from r = .25 (satisfaction with leisure time and satisfaction with one's individual domain) to r = .91 (satisfaction with one's primary activity in general and satisfaction with the tasks at one's primary activity). Within-person correlations were also consistently positive—on days that people reported higher satisfaction than usual with one aspect, they also tended to report higher satisfaction than usual with other aspects—but consistently smaller, ranging from r = .09 (satisfaction with leisure time and satisfaction with social contacts at one's primary activity) to r = .63 (satisfaction with one's primary activity in general and satisfaction with the tasks at one's primary activity).

Explicit Assessments of Role of Life Domains (Initial and Final Questionnaire; RQ 5)

In the initial questionnaire, participants reported how relevant they considered each life domain which applied to them in three different ways. We subsume these measures under the label of "explicit assessments" as they reflect information that is available and reported by the participant (in contrast, the individual effects of domain satisfaction on general satisfaction that we estimated may be considered an implicit assessment of the role of life domains). In the present article, we will only consider importance rating and perceived influence.[3] The assessment was repeated in the final questionnaire, which allowed us to investigate the retest properties of the included measures. However, when investigating whether these explicit assessments correlate with the individual-level effects of domain satisfaction on general satisfaction, we relied only on the initial assessment as they can be considered "pretreatment" variables (Montgomery et al., 2018) and thus avoid some interpretational issues (e.g., reflecting on domain satisfaction throughout the study may affect the perceived importance).

Importance Ratings. Participants were asked to indicate the importance of each applicable life domain on a five-point scale ranging from *not important at all* to *very important*.

Perceived Influence. Participants were asked to consider how strongly each relevant domain influences their happiness. They reported their guess on a five-point scale ranging from *no influence* to *strong influence*.[4]

Reliability of Explicit Assessments of Role of Life Domains

Previous studies may have failed to detect reliable evidence for importance weighting due to a lack of reliability of single-item importance ratings—an issue that was already raised in the seminal study on importance weighting (Campbell et al., 1976, pp. 87–88). We thus calculated test-retest Pearson correlation coefficients between the initial and final questionnaire for importance ratings and perceived influence. Test-retest correlations for the importance assessments ranged between .46 and .74 for importance ratings, and between .41 and .68 for perceived

3 The ranking consisted of participants creating a Top 5 list of domains that mattered most to them, which they then ranked from 1 to 5. This procedure makes it impossible to assign the same level of importance to multiple domains at once and can thus potentially circumvent the lack of variability in importance ratings observed in other studies (Hsieh, 2003, 2012). However, the ranking comes with some analytic complications (e.g., some domains rarely making it into the Top 5) and as our focus shifted away from importance weighting, we decided against analysing it for the present article.

4 They also reported their confidence in their ratings on a five-point scale ranging from *very uncertain* to *very certain*; we did not analyze these confidence items for the purpose of the present study.

influence (see Table 3). This suggests that the retest reliabilities of the single-item domain assessments may be at least in the medium range, with some variability between domains. For example, test-retest correlations were higher for the domain family than for the domain leisure use. Over the course of the study, the actual importance of the life domains may have changed (maybe in part because of heightened attention due to the repeated diary questionnaire); thus, retest correlations can be taken as a lower bound estimate of the reliability of the single items.

Importance ratings and perceived influence were highly correlated with correlations ranging from .52 to .77 in the initial questionnaire. We thus averaged them to arrive at a combined rating that is potentially more reliable (all resulting αs > .70, see Table 3). This combined rating had test-retest correlations ranging from .53 to .79, with all values exceeding the corresponding test-retest correlations for the single-item ratings.

Big Five (Initial Questionnaire)

At the end of the initial questionnaire, participants filled out a German translation of the BFI-2-S (Danner et al., 2016; Rammstedt et al., 2020; Soto & John, 2017),[5] which contains six items for each of the Big Five traits (neuroticism, extraversion, conscientiousness, openness, agreeableness). Participants answered all items on a five-point scale (1 = *do not agree at all* to 5 = *fully agree*). The resulting internal consistencies of the scales were satisfactory ($\alpha_{Neuroticism}$ = .81, $\alpha_{Extraversion}$ = .79, $\alpha_{Conscientiousness}$ = .76, $\alpha_{Agreeableness}$ = .70, $\alpha_{Openness}$ = .70)

Analysis Strategy

Analyses were conducted in two steps. In the first step, we estimated participants' individual-level effects as within-person slopes in multilevel models, predicting general satisfaction from domain satisfaction at the same point in time (RQs 1-2). In the second step, to explore potential explanations, we correlated these individual-level effects with various other variables (RQs 3-6). We chose this two-step approach because (1) substantively, it fits the general approach of our study, as we aim to generate and describe measures of interindividual differences and then explore their associations with various third variables, and (2) pragmatically, we ran into problems when trying to implement a one-step approach given the complexities of the multilevel models with nested predictors and indirect effects.

We use Bayesian 95% credible intervals (CIs) to evaluate the strength of the statistical evidence/precision of the estimates. Much like frequentist confidence intervals, these intervals capture uncertainty in the point estimates (Albers et al., 2018). They contain the unobserved parameter with a probability of 95%, taking into account the information provided by the prior.

Estimating Participants' Individual Effects of Life Domains on General Satisfaction (RQ 1)

We ran a Bayesian multivariate multilevel model to estimate participants' individual-level effects. Figure 2 provides a conceptual overview of the analysis. This analysis was conducted with the R package *brms* (Bürkner, 2017) using default priors (flat priors for population-level effects, weakly informative priors for variances). To remove between-subjects differences in the predictors, we person-mean (i.e., within-subject) centered all domain satisfaction ratings (see section on causal identification below). The multilevel models included random intercepts (which capture differences between individuals in the outcome) and random slopes for all predictors (which capture differences between individuals in the coefficients of interest; that is, the relevant between-person variation in the link between domain satisfaction on general satisfaction). Not all domains applied to all participants and thus, not all predictor variables were available for all participants. To deal with this without excluding participants, we included the affected domain ratings (see Figure 2) as nested predictors, which ensures that the corresponding coefficients are informed only by the participants to whom the domain applied. Technically, this is achieved by including two predictors per affected domain: a binary indicator which reflects whether the domain applied to the participant, and the product of this binary indicator with the respective domain satisfaction rating. For participants to whom the domain did not apply, the product automatically equals zero and thus, their (non-existent) values on the non-applicable domain do not inform the analysis.[6]

The analysis was multivariate because it simultaneously modeled three outcome variables: general satisfaction, satisfaction with one's social life, and general satisfaction with one's primary activity. General satisfaction was the ultimate outcome we were interested in; it was predicted by all domain satisfaction ratings included in the study provided at the same time point (Figure 2). The other two outcomes were not of interest per se but rather intermediate outcomes (which in turn predict general satisfaction); modeling them simultaneously allowed us to include all domains in a single model while still being able to recover all effects of interest (see section on causal identification). Satisfaction with one's social life was predicted by satisfaction with the domains family, friends, relationship, and primary activity (PA): social contacts. General satisfaction with one's primary activity was predicted by satisfaction with PA: social contacts, PA: tasks, and PA: performance.

5 Our Big Five items slightly deviate from the official German translation since we considered some of the phrasings not optimal or too different from the English original (see codebook provided on the OSF, https://osf.io/x8j7r/, for item text).

6 For the model to run, these non-existent values need to be set to some arbitrary value, the choice of which does not affect the results.

Table 3. Item Characteristics of the Explicit Assessments of Life Domains

		Importance rating[a]		Perceived influence[a]		Combined rating[a, b]			Test-Retest Correlations [95% CI]			
Life Domain	*N*	*M*	*SD*	*M*	*SD*	*M*	*SD*	$\alpha_{Combined}$[a, b] [95% CI]	N_{Retest}	Importance rating	Perceived influence	Combined rating
Health	439	4.4	0.8	4.3	0.8	4.4	0.7	.70 [.64; .76]	361	.62 [.55; .68]	.54 [.46; .61]	.67 [.61; .72]
Social life	439	3.9	0.9	3.8	0.9	3.9	0.8	.76 [.71; .81]	361	.59 [.52; .65]	.52 [.44; .59]	.65 [.58; .70]
Family	439	4.1	1.0	4.0	1.0	4.1	0.9	.86 [.83; .88]	361	.74 [.69; .79]	.68 [.62; .73]	.79 [.74; .82]
Friends	439	4.2	0.9	4.0	0.9	4.1	0.8	.82 [.78; .86]	361	.69 [.63; .74]	.66 [.60; .72]	.77 [.72; .81]
Looks	439	3.3	0.9	3.2	1.0	3.3	0.9	.87 [.84; .90]	361	.69 [.63; .74]	.62 [.55; .68]	.73 [.68; .77]
Leisure time	439	4.1	0.9	4.0	0.9	4.1	0.8	.80 [.76; .84]	361	.59 [.52; .65]	.49 [.41; .57]	.63 [.56; .69]
Leisure use	439	4.0	0.9	4.0	0.9	4.0	0.8	.82 [.79; .86]	361	.46 [.38; .54]	.41 [.32; .49]	.53 [.45; .60]
Relationship	248	4.5	0.7	4.4	0.7	4.5	0.6	.75 [.70; .80]	194	.69 [.60; .75]	.52 [.41; .62]	.74 [.67; .79]
Exercise	276	3.7	0.9	3.7	0.9	3.7	0.9	.84 [.81; .87]	222	.67 [.59; .74]	.62 [.53; .69]	.69 [.61; .76]
Primary activity												
General	369	4.1	0.9	4.1	0.8	4.1	0.7	.68 [.62; .75]	289	.71 [.64; .76]	.51 [.42; .59]	.69 [.62; .74]
Tasks	369	4.1	0.8	4.0	0.8	4.0	0.8	.75 [.70; .80]	289	.59 [.51; .66]	.50 [.40; .58]	.65 [.58; .71]
Performance	369	4.1	0.8	4.0	0.9	4.0	0.8	.78 [.73; .82]	289	.56 [.47; .63]	.54 [.45; .61]	.65 [.58; .71]
Social contacts	369	3.6	0.9	3.5	0.9	3.6	0.9	.85 [.82; .88]	289	.65 [.58; .71]	.58 [.50; .65]	.68 [.62, .74]
Individual domain	162	4.5	0.6	4.5	0.6	4.4	0.6	. 84 [.81; .87]	131	.53 [.40; .64]	.60 [.48; .70]	.65 [.54; .74]

Note. Importance ratings on a scale from 1 (*not important at all*) to 5 (*very important*); perceived influence from 1 (*no influence*) to 5 (*strong influence*). 95% CI refers to the 95% confidence interval.

[a] In the initial questionnaire.

[b] Combined rating is an average of the importance rating and the perceived influence. As the combined rating consists of two items, the reported Cronbach's α values can be transformed into the corresponding Pearson correlations: $r = \alpha/(2-\alpha)$.

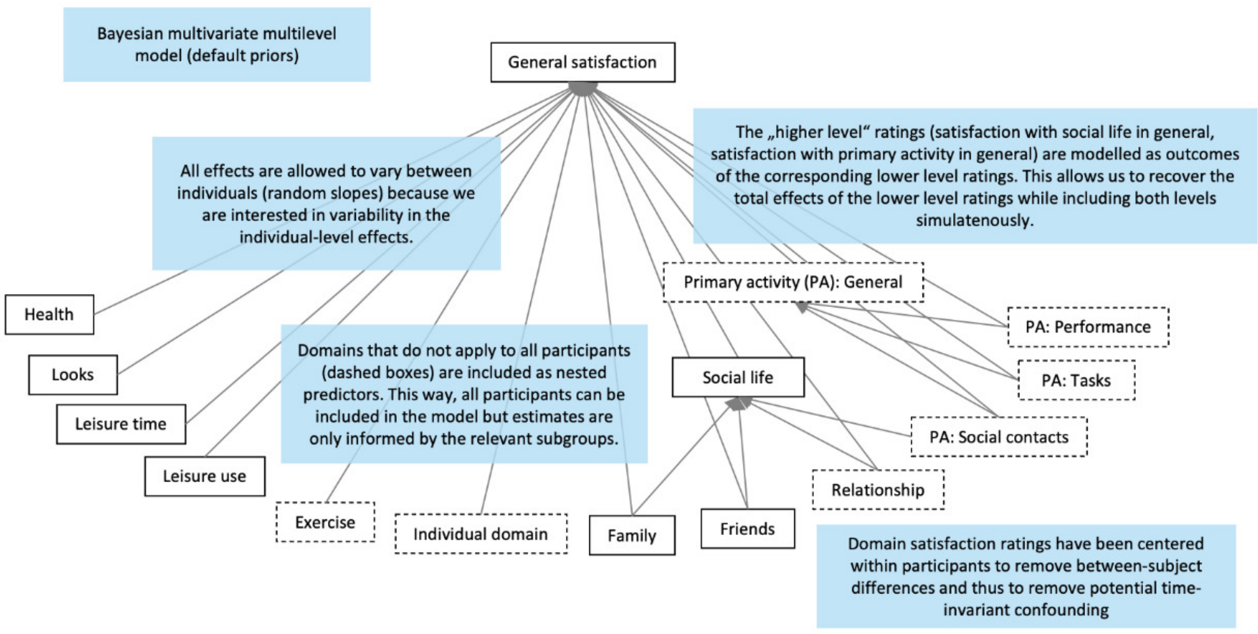

Figure 2. Overview of the Multilevel Model Used to Estimate Participants' Individual Effects

After estimating the model, for each person and each domain, we extracted the estimates of the effects of domain satisfaction on general life satisfaction, including the standard deviation of each person-specific coefficient's posterior samples. Posterior samples are drawn from the posterior distribution of each estimated parameter and inform us about the probability of each parameter value given the data and our priors (here the non-/weakly-informative *brms* default prior). Conceptually, the standard deviation of the posterior samples reflects the uncertainty of the estimate; in frequentist statistics, this would correspond to the standard error. This allows us to propagate uncertainty across analyses—for example, if we are very uncertain about the effect of health satisfaction on general satisfaction for a particular individual (for instance, because they contributed few days or exhibited little variability in health satisfaction), that individual's effect would be down-weighted when we correlate the individual-level effects with third variables to explore potential explanations.

Considerations Regarding the Precision of the Estimated Individual Effects (RQ 2)

We quantified the precision of the estimated individual effects by investigating credible intervals for the individual estimates, and by calculating a reliability coefficient that takes into account the variance within participants relative to the variance between participants:

$$Reliability = 1 - \frac{\left(Standard\ error\ of\ \hat{b}\right)^2}{\left(SD\ of\ \hat{b}\ across\ individuals\right)^2}$$

with the squared standard error calculated from the posterior distribution of the individual-level effects (by squaring the standard error of each individual-level effect estimate and then averaging across individuals). This formula is equivalent to the standard formula for reliability from classical test theory, with the terms in the numerator and the denominator shrunk by the reliability,

$$\begin{aligned} Reliability &= 1 - \frac{(Standard\ error\ of\ b)^2}{(SD\ of\ b\ across\ individuals)^2} \\ &= 1 - \frac{Reliability * \left(Standard\ error\ of\ \hat{b}\right)^2}{Reliability * \left(SD\ of\ \hat{b}\ across\ individuals\right)^2} \\ &= 1 - \frac{\left(Standard\ error\ of\ \hat{b}\right)^2}{\left(SD\ of\ \hat{b}\ across\ individuals\right)^2}. \end{aligned}$$

In principle, the precision with which the individual effects can be estimated depends on multiple factors. First, it will depend on the amount of stable interindividual differences. If there are only small interindividual differences, these will be estimated rather imprecisely. Likewise, if there are large interindividual differences but these are not stable over time (e.g., for some people, on some days health satisfaction may have a large effect; on other days, health satisfaction may be less important as other, more salient, events occur), again estimates will be imprecise. Second, it will depend on the intercorrelations between the different domain satisfaction scores—this is because our causal identification strategy (next section) requires that all scores are included simultaneously; estimates become more imprecise when predictors are correlated. Third, the precision also depends on the design, in particular the number of observations per individual (here, $5 \leqslant n \leqslant 15$). In analogy to standard latent variable models, one can think of the individual effect as the latent factor and the individual diary entries as items; all else being equal, the reliability will increase as the number of items (the number of diary entries per individual) increases.

Causal Identification Assumptions of the Individual Effects

The coefficients from the multilevel model only reflect the causal effects of domain satisfaction on general satisfaction when certain assumptions about a lack of unobserved confounding are met, and these assumptions in turn depend on the specifics of the estimated model. First, we could be worried about confounders that operate "between people" such as gender (i.e., time-invariant confounders). However, domain satisfaction scores were within-subject centered, which effectively removes the influence of such confounders (Rohrer & Murayama, 2023). This leaves us to worry about confounders that operate "within people," that is, variables that affect both domain satisfaction and general satisfaction, and that vary within individuals over the course of the study (i.e., time-varying confounders).

We assume a model in which general life satisfaction is the outcome of various domain satisfaction ratings, and these domain satisfaction ratings are in turn caused by "bottom-up influences," which subsume all the things happening in our lives. Consider the scenario on the left side of Figure 3, Panel A. Here, job satisfaction and health satisfaction share a common cause, back pain that a participant may experience on some days (reducing their health satisfaction but also their job satisfaction as their work becomes more straining). If we only looked at the bivariate (within-person) association between health satisfaction and general satisfaction, this causal structure would be a problem because there would be a so-called open backdoor-path, Health satisfaction ← Current back pain → Job satisfaction → General satisfaction. Such a path introduces non-causal associations between the variables of interest. However, if job satisfaction is included in the model (and thus "statistically controlled for"), this path is blocked and no longer leads to non-causal associations (for more explanation regarding the logic of third-variable control, see Rohrer, 2018; Wysocki et al., 2022). In the presence of measurement error, some non-causal associations may remain (Westfall & Yarkoni, 2016) but the resulting bias would be reduced. For this reason, we include *all* measured domains when estimating the effects of domain satisfaction; it plausibly reduces the confounding influences of life circumstances that affect multiple domains at once.

A problematic scenario arises if an included domain shares common causes with a domain that we forgot to include, such as depicted on the right side of Figure 3, Panel A. Here, a participant has an argument with their highly religious father which lowers both their family satisfaction (which we did include) and their faith satisfaction (which we did not include, but which has been included in other studies, e.g., Rohrer & Schmukle, 2018). This confounds family satisfaction with faith satisfaction, and since we did not measure faith satisfaction, we cannot close this non-causal path by including the variable in the model. Thus, we will wrongly attribute part of the effect of faith satisfaction to family satisfaction. For our analysis, this means that a causal interpretation of any estimated individual-level effect of domain satisfaction on general satisfaction hinges on whether or not we succeeded in including all life domains that are relevant to said individual. Thus, that we allowed participants to add a "custom" life domain potentially makes causal identification more plausible.

A less transparent causal inference concern arises because we included two domain satisfaction ratings (i.e., satisfaction with social life and general satisfaction with primary activity) that plausibly aggregate other domain satisfaction ratings included (e.g., family, relationship; primary activity: tasks, primary activity: performance). Conceptually, these two higher-level domains may mediate the effects of the corresponding lower-level domains. If we simply run a model predicting general satisfaction from all domain ratings, including the higher *and* the lower levels, the coefficients associated with the lower level change their meaning: for example, the coefficient associated with family satisfaction would be the predicted change in general satisfaction for a change in family satisfaction, holding social life satisfaction constant. One may hope that this coefficient reflects the direct effect of family satisfaction, excluding any part mediated via social life satisfaction. However, this only holds if we successfully included all lower-level domains that causally affect social life satisfaction.

Consider the scenario in Figure 3, Panel B. Here, we failed to measure congregation satisfaction (which may be relevant for religious participants) which in turn affect social life satisfaction. Our analysis conditions on social life satisfaction. This variable is a so-called collider (Rohrer, 2018); conditioning on it will introduce a spurious association between its causes (family, relationship, and unmeasured congregation satisfaction) which is negative in the simple linear, additive case. Intuitively, if somebody reports the same level of social life satisfaction as usual for them, but much lower satisfaction with their family and their relationship, we know that something else must be going on—some other social aspect of their life must be going well (e.g., congregational life). Thus, we may introduce a spurious association between an included domain of interest and general satisfaction via an unobserved domain (Family satisfaction ↔ Congregation satisfaction → General satisfaction) which biases the estimate of the (supposedly) direct effect. To avoid that this type of collider bias leads to mistaken conclusions, we do not interpret the coefficients of lower-level domains from models that include the higher-level domains. Instead, we derive the *total* effects of the lower-level domains, including any parts mediated via the higher-level domains. To do so, we additionally explicitly model the higher-level domains, which allows us

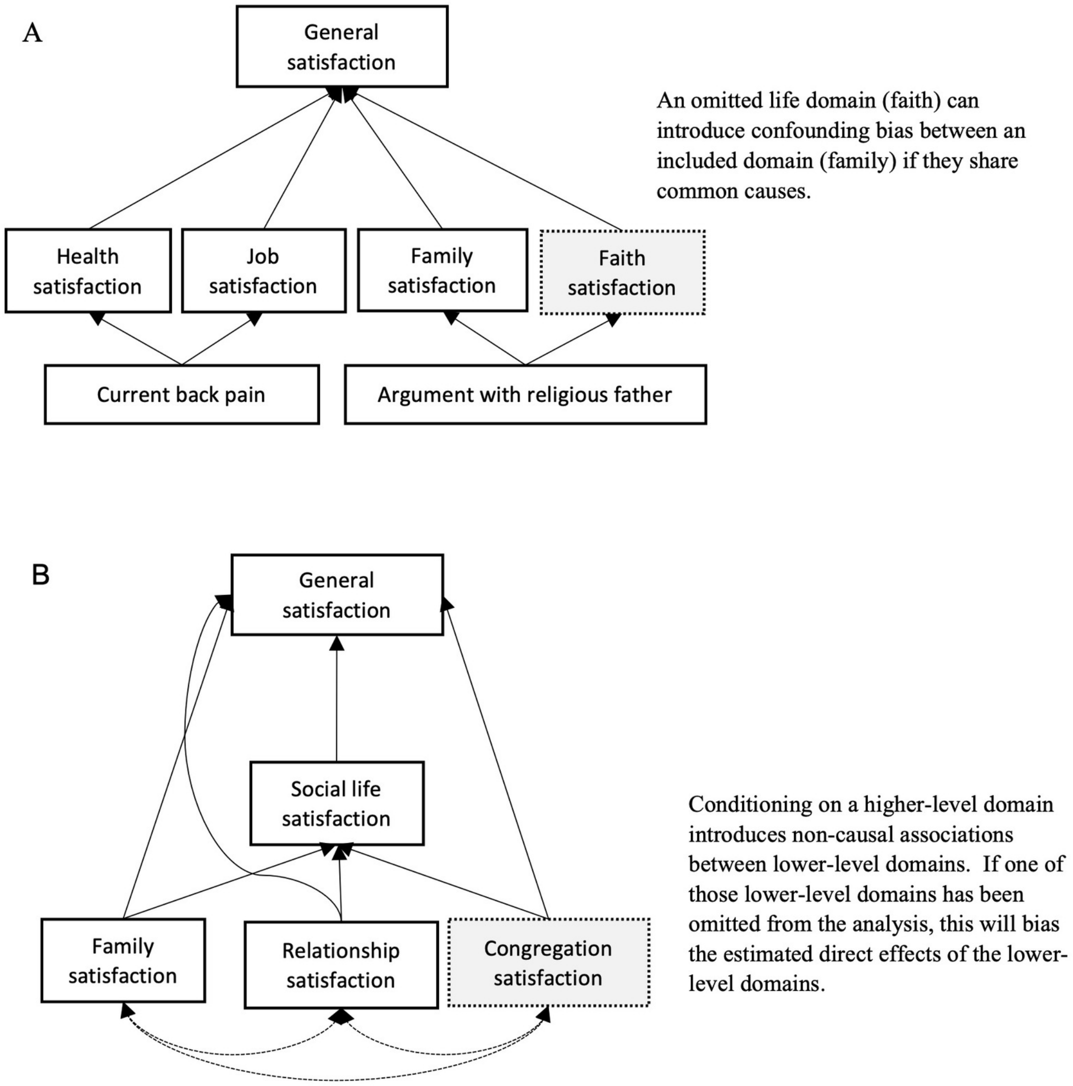

Figure 3. Model in Which Life Circumstances Affect Domain Satisfaction, Which in Turn Affects General Satisfaction

to, in the terms of mediation analysis, add up direct and indirect effects, which in turn cancels out the type of collider bias described above.[7]

Further Analysis of the Estimated Individual Effects

After extraction of the individual-level effects of domain satisfaction on general life satisfaction, we analyzed these estimates further. For each domain, we ran Bayesian regression analyses predicting the individual-level effects of the preceding analysis from various (between-subject) predictors listed below. The *brms* package makes it possible to incorporate information about uncertainty in the outcome via the *se*() argument.[8] We used this argument and passed on the standard deviation of the posterior samples of the effect of interest.

In general, it is not justified to give the resulting associations between individual-level effects and interindividual differences a causal interpretation. If we observed that participants who say that their family is very important to them also experienced larger effects of family satisfaction, it would be tempting to conclude that those effects

7 Note that in the scenario we just described (Figure 3, Panel B), deriving the total effects of the lower-level domains only means that *their* effect estimates are not confounded by the omitted life domain. The effects of the higher-level domain on general satisfaction would still be confounded (congregation satisfaction is an unobserved confounder between social life satisfaction and general satisfaction).

8 For example, this would be standard practice in meta-analyses as the effect estimates from the individual studies are also to some degree uncertain.

are larger because their family is important to them (importance of domain → effect of domain on general satisfaction). But this association could also arise from reverse causality (participants observe their family has an outsized impact on their general life satisfaction and thus conclude it is important; effect of domain on general satisfaction → importance of domain), and, much more importantly, confounding factors. For example, maybe participants who have children rate their family as more important, and family satisfaction becomes a much more important determinant of general satisfaction once children are around (importance of domain ← children → effect of domain on general satisfaction).

In contrast, associations between effects and, for example, gender, may appear less problematic from a causal inference perspective. Both reverse causality and confounding seem implausible. But this only holds at the population level. Selection into the study sample can induce collider bias between all variables that affect whether people participated in the study in the first place (Rohrer, 2018). For example, consider a situation in which women are more likely to participate and people who are more socially engaged are more likely to participate (Figure 4). Social engagement in turn affects the effect of social satisfaction on general satisfaction. Then, within the study (conditional on study participation, the collider), a spurious association between gender and social engagement is induced, which can in turn lead to a spurious association between gender and the effect of social satisfaction on general satisfaction.

Correlating Effects with Participant's Average and Variability of Domain Satisfaction (RQ 3). Interindividual differences in the effects may arise "mechanistically" from the distribution of domain satisfaction—for example, maybe participants with lower average satisfaction experience larger effects, or maybe participants with less variability in domain satisfaction experience smaller effects. For each domain, we regressed the individual-level effect of domain satisfaction onto both the person-specific mean of the domain satisfaction rating and the person-specific standard deviation of the domain satisfaction rating, for example, effect of health satisfaction on general satisfaction ~ individual mean(health satisfaction) + individual sd(health satisfaction). We had no particular expectations considering associations between the effects and the mean of domain satisfaction. Considering associations between the effects and the standard deviation of domain satisfaction, intuitively one may suspect that if an individual experiences more variability in domain satisfaction, then said domain should have a bigger effect. However, this intuition only applies to standardized effect size metrics, at which we would arrive if we within-subject standardized effect estimates (by multiplying the unstandardized estimate with the person-specific standard deviation of domain satisfaction and divide by the person-specific standard deviation of general satisfaction). Without such standardization, whether variability correlates with effect estimates remains an empirical question.[9]

Correlating Effects with Gender and Age (RQ 4). For each domain, we also regressed the estimated effect on (a) gender and (b) age. For the gender analyses, we excluded a small number of participants (1%) who had not identified as either a woman or a man and then regressed the individual-level effects on a binary indicator of whether the participant was a woman.

Correlating Effects with Explicit Assessments of Role of Domain (RQ 5). For each domain, we regressed the estimated effect of the domain on the combined rating (mean of importance rating and perceived influence) of the domain (e.g., effect of health satisfaction ~ combined rating of role of health).

Correlating Effects with Big Five (RQ 6). For each domain, we regressed the estimated effect on all Big Five personality traits in a single model (effect of domain on general satisfaction ~ neuroticism + extraversion + conscientiousness + openness + agreeableness).

Software

Analyses were conducted in *R* (R Core Team, 2022) with the help of *RStudio* (RStudio Team, 2020), using the package *brms* (Bürkner, 2017). Additionally, we used the packages *data.table* (Dowle & Srinivasan, 2021), *formr* (Arslan et al., 2020), *ggplot2* (Wickham, 2016), *knitr* (Xie, 2022), and *psych* (Revelle, 2017).

Results

Effects of Domain Satisfaction on General Satisfaction

Averages of the Individual-Level Effects

Estimates from the multilevel model indicated positive average effects of domain satisfaction on general satisfaction for all life domains. Table 4 displays the estimates, including a standardized parameter estimates for which the average effect was multiplied with the standard deviation of (within-subject centered) domain satisfaction and divided by the standard deviation of the (within-subject centered) outcome scale (i.e., 1.29). The largest average effect was observed for satisfaction with leisure use (b = 0.16 points on the adapted SWLS per point of leisure use satisfaction, or b_{std} = 0.25 standard deviations of the outcome per standard

9 Regardless of any associations between the effects of domain satisfaction and the mean and standard deviation of domain satisfaction, if a participant's satisfaction with a domain barely varies (low person-specific standard deviation), their individual-level effect of domain satisfaction on general satisfaction will be estimated imprecisely (high standard deviation of the posterior distribution of the estimate). We indeed observed this pattern in our data with medium to strong negative associations (correlations between person-specific standard deviation of domain satisfaction and standard deviation of the posterior of the effect estimate between −.31 and −.83). As described above, further analyses took into account the precision with which individual-level effects were estimated.

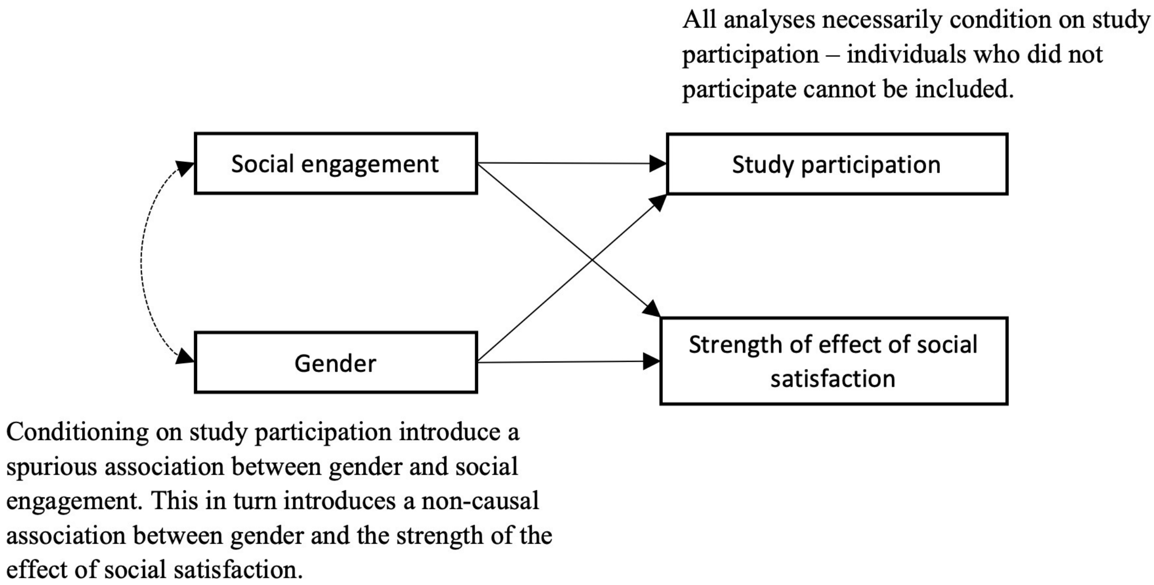

Figure 4. Selective Study Participation Can Introduce Non-Causal Associations Between Interindividual Differences and Individual-Level Effects

Table 4. Results from the Multilevel Model Estimating the Effects of Domain Satisfaction on General Satisfaction

	Average effect		Effect variability		Standardized
Life Domain	*b*	95% CI	*SD*	95% CI	average effect[a]
Health	0.11	[0.09; 0.13]	0.12	[0.10; 0.14]	.15
Social life	0.15	[0.12; 0.18]	0.14	[0.10; 0.17]	.15
Family	0.07	[0.04; 0.09]	0.12	[0.09; 0.15]	.06
Friends	0.11	[0.08; 0.14]	0.12	[0.08; 0.15]	.11
Looks	0.07	[0.04; 0.10]	0.15	[0.11; 0.19]	.06
Leisure time	0.10	[0.08; 0.12]	0.11	[0.09; 0.14]	.14
Leisure use	0.19	[0.16; 0.21]	0.14	[0.11; 0.16]	.25
Relationship	0.16	[0.12; 0.19]	0.13	[0.10; 0.16]	.17
Exercise	0.01	[−0.01; 0.04]	0.09	[0.06; 0.12]	.02
Primary activity					
General	0.09	[0.06; 0.12]	0.14	[0.09; 0.18]	.09
Tasks	0.10	[0.07; 0.12]	0.11	[0.08; 0.14]	.10
Performance	0.04	[0.02; 0.07]	0.08	[0.04; 0.12]	.05
Social contacts	0.06	[0.04; 0.09]	0.11	[0.07; 0.14]	.07
Individual domain	0.09	[0.06; 0.13]	0.05	[0.00; 0.11]	.11

Note. For life domains nested within higher-level domains estimates refer to total effects and are thus not conditional on the higher-level domain. 95% CI refers to the 95% credible interval.
[a]Average effects were standardized by multiplying the estimate with the standard deviation of (within-subject centered) domain satisfaction and dividing by the standard deviation of the (within-subject centered) adapted SWLS.

deviation of the predictor), followed by relationship satisfaction (b = 0.16, b_{std} = 0.17), social life satisfaction (b = 0.15, b_{std} = 0.15) and health satisfaction (b = 0.11, b_{std} = 0.15). The smallest average effect was observed for exercise satisfaction (b = 0.01, b_{std} = 0.02), for which the 95% credible interval (CI) also included zero.

Our main analyses treated all primary activities as equal, but one could plausibly expect that satisfaction with one's full-time job plays a different role than satisfaction with one's studies. We thus additionally investigated whether the estimated effects of satisfaction with one's primary activity varied between the three biggest groups (college, n = 134; full-time work, n = 160; part-time work, n = 57). Results suggested that the average effect of overall satisfaction with one's primary activity (Primary activity: General) *could* be bigger among those who work full-time compared to college students (b_{diff} = 0.01, 95% CI: [-0.01, 0.04]), but the evidence for differences between those groups was weak in general, possibly due to the low sizes of the groups being compared. More detailed results comparing the different primary activities are reported in the online supplement provided on the OSF, https://osf.io/x8j7r/.

Interindividual Variability in the Individual-Level Effects (RQ 1)

Furthermore, across participants, the individual-level effect estimates showed considerable variability in almost all domains (Figure 5). This variability was particularly high for domains for which average effects were also high ($SD_{Leisure\ use}$ = 0.14, $SD_{Social\ life}$ = 0.14, $SD_{Relationship}$ = 0.13). Considering the two domains of which we had hoped they were not highly important to everyone, satisfaction with one's looks indeed had the highest variability in effects (SD = 0.15), whereas exercise satisfaction showed relatively smaller, albeit non-zero variability (SD = 0.09). The smallest variability was observed for satisfaction with the individual domain that participants could add to the questionnaire ($SD_{Own\ domain}$ = 0.05). In hindsight, we believe that this lack of variability arises "by design": participants were asked whether anything important was missing among the proposed domains which sets a lower bound for the individual-level effects to be expected here; at the same time, the included domains already covered major aspects of life which sets an upper bound. Given this lack of variability, we excluded participants' own domains from subsequent analysis; there are barely any interindividual differences in effects to be explained here and furthermore, associations would be hard to interpret given that this domain captures different aspects of life for different people. Despite the overall noticeable variability in individual-level effects of domain satisfaction, for almost all domains, the point estimate of the individual-level effect was positive for the large majority of participants (> 90%), the two exceptions being satisfaction with one's looks (82% of effects > 0) and exercise satisfaction (70% of effects > 0).

Estimates of the Individual-Level Effects are Rather Unreliable (RQ 2)

The previously reported results give us confidence that the effects of domain satisfaction on general satisfaction vary between individuals, with at least some degree of stable interindividual differences. If the observed differences in the effects were resulting from chance or measurement error alone, the *SD*s of the respective random effects would be close to zero as the model takes into account chance variation and estimates the distribution of the "true" underlying effects. However, just because we know that there are interindividual differences does not mean that we can measure them reliably. The point estimates of individual-level effects were associated with large degrees of uncertainty, which Figure 6 illustrates for four domains (the corresponding figure including all domains can be found in an online supplement provided on the OSF, https://osf.io/x8j7r/). For almost all individuals, we cannot rule out considerably larger or smaller effects; for many, zero or even negative effects are compatible with the data. This is also reflected in reliability estimates, for example, considering the four domains depicted in Figure 6: $Reliability_{Health}$ = .33, $Reliability_{Social\ life}$ = .28, $Reliability_{Looks}$ = .27, $Reliability_{Leisure\ use}$ = .35. This uncertainty necessarily reduces the precision with which we can estimate correlations between the individual-level effects of domain satisfaction and other variables.

Correlations with M and SD of Domain Satisfaction (RQ 3)

Can some of the variability in the effects be predicted from an individual's average satisfaction in the respective domain, or from their variability in satisfaction in the respective domain? Our regression analyses suggested that this *may* be the case for at least two domains, health and relationship (Figure 7, upper panel). Participants who reported lower average levels of health satisfaction experienced larger effects of health satisfaction on general satisfaction (b = −0.014). Likewise, participants who reported lower average levels of relationship satisfaction experienced larger effects of relationship satisfaction (b = −0.017). There was no evidence for any associations between variability in domain satisfaction ratings and the effects of domain satisfaction (Figure 7, lower panel).

Correlations with Gender and Age (RQ 4)

There were no noticeable gender differences in the individual-level effects of domain satisfaction on general satisfaction (*b*s between -0.009 and 0.013, all 95% CIs included 0; more detailed results in the online supplement provided on the OSF). Considering age, analyses indicated that the effect of family satisfaction may be larger among older participants (b = 0.01 per year of age, 95% CI [0.001; 0.020]; more detailed results in the online supplement provided on the OSF). Figure 8 visualizes this linear association alongside a locally smoothed curve. For all other domains, the coefficient was smaller, and 95% CIs included 0. This of course may also be a result of the imprecision with which the individual-level effects were estimated.

Correlations with the Explicit Assessments of the Role of Domains (RQ 5)

Can the individual-level effects of domain satisfaction on general satisfaction be predicted from participants' ratings (i.e., average of importance and perceived influence) of said domain? In general, the ratings only explained very little of the interindividual differences in effects (Figure 9). At the same time, at least two of the domains looked like there may be some evidence for a positive association between ratings and effects (looks, leisure time; Figure 9). Again, the uncertainty in all estimates means that we cannot rule out small positive (or even negative) associations for other domains.

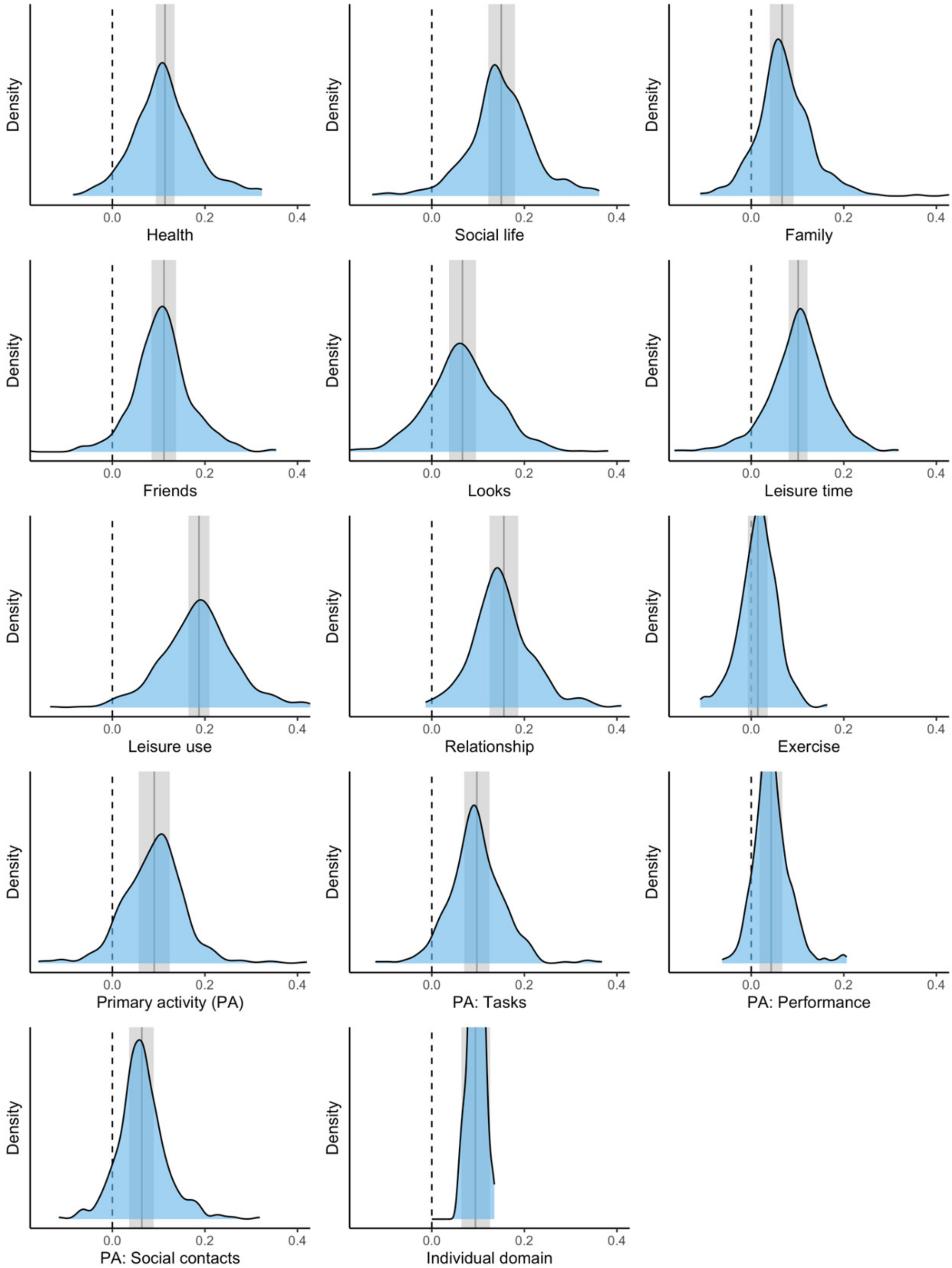

Figure 5. Distribution of the Point Estimates of the Individual-Level Effects of Domain Satisfaction on General Satisfaction

Note. Estimates of the individual-level effects of satisfaction with the respective life domain on general satisfaction from our multilevel model ($N_{Individuals}$ = 439, $N_{Diary\ questionnaires}$ = 6,071). Effects are expressed in the original units; points on the general satisfaction scale (which ranges from 1 to 7) per point on the domain satisfaction scale (which ranges from 1 to 7). The gray vertical line indicates the average effect, and the shaded area indicates the 95% credible interval around the average effect.

Correlations with the Big Five (RQ 6)

Lastly, for each domain, we tried to predict the individual-level effects from participants' Big Five traits. In total, the Big Five only explained modest amounts of variance of the effects of interest with R^2 ranging from 3% to 9%, with the corresponding multiple correlation coefficients R between .17 and .31 (Figure 10). When looking at the individual coefficients for each trait and all domains, no particular

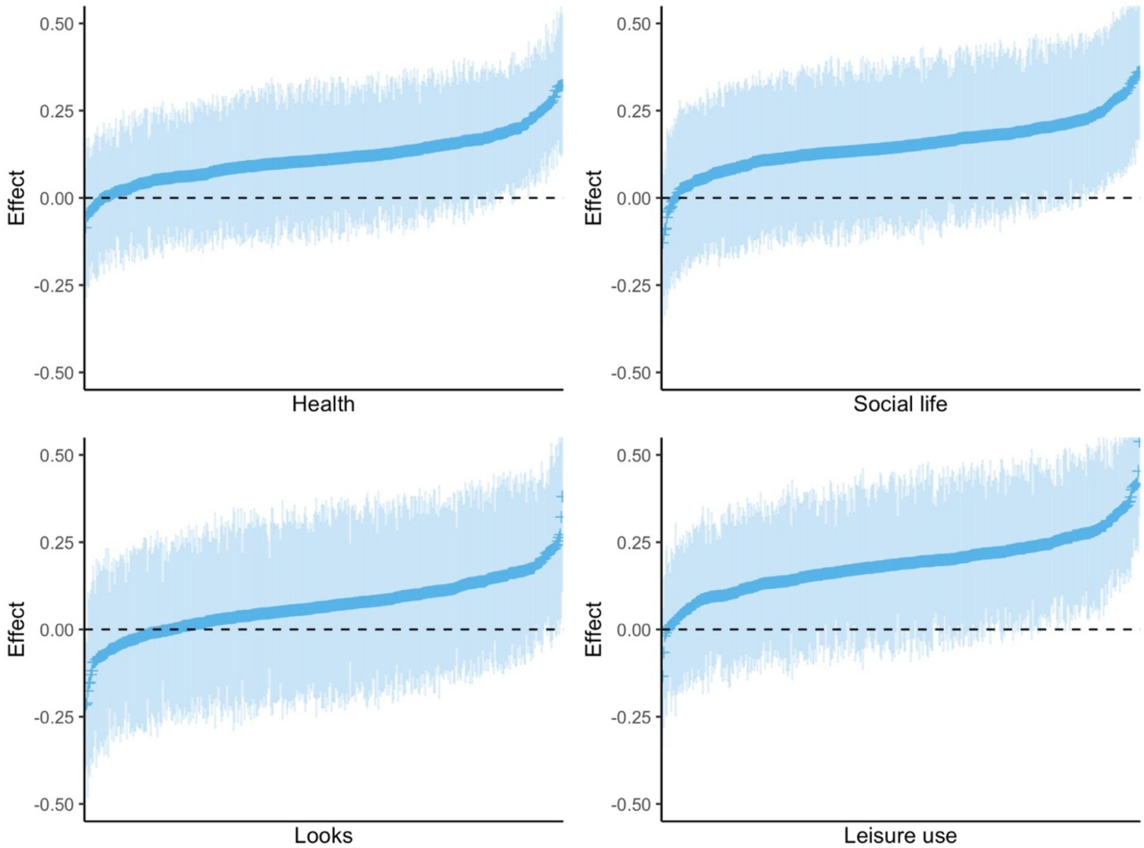

Figure 6. Estimated Individual-Level Effects of Domain Satisfaction, Sorted by Magnitude, for Four Domains,

Note. Estimates of the individual-level effects of satisfaction with the respective life domain on general satisfaction from our multilevel model ($N_{Individuals}$ = 439, $N_{Diary\ questionnaires}$ = 6,071). Effects are ordered by magnitude and expressed in the original units; points on the general satisfaction scale (which ranges from 1 to 7) per point on the domain satisfaction scale (which ranges from 1 to 7). Blue vertical lines are 95% credible intervals.

overarching pattern (such as "high neuroticism always predicts stronger effects") stood out to us.

The lower panels of Figure 10 illustrate the results for four domains that we deemed of interest because the Big Five explained relatively more variance (relationship and exercise satisfaction) or because the domain showed relatively higher effect variability (looks and leisure use satisfaction). Considering relationship satisfaction, results suggested that participants who scored higher on neuroticism and possibly agreeableness experienced larger effects of relationship satisfaction on general satisfaction. Considering exercise satisfaction, maybe participants who scored high on extraversion experienced larger effects. Considering the effects of satisfaction with the way one looks, here again, individuals with higher neuroticism may have experienced larger effects. Lastly, considering the effects of how satisfied one was with their leisure use, these seemed to be smaller among more extraverted participants. All of the estimates came with considerable uncertainty, so that we cannot draw any firm conclusions. The online supplement contains plots for all domains.

General Discussion

Average Effects of Domain Satisfaction on General Satisfaction

In this study, we aimed to identify the individual-level causal effects of satisfaction with various life domains on general life satisfaction with the help of longitudinal diary data. Overall, the results indicated that, on average across participants, satisfaction with all domains had positive effects on general life satisfaction. If we compare the average effects across domains, the results appear overall sensible, with smaller effects for aspects that are likely less focal to participants' lives (such as the way they look, physical exercise, social contacts at work) compared to more focal ones (such as health, social life in general, romantic relationships). This may appear trivial; however, previous studies do not readily allow for the comparison of effects across domains given that they either do not plausibly identify effects (cross-sectional studies on importance weighting) or only consider individual factors in isolation (longitudinal studies).

However, we did not anticipate that ratings of satisfaction with how one had used one's leisure would have the overall largest effects. It is possible that the estimated effects for this domain were larger because it served as a

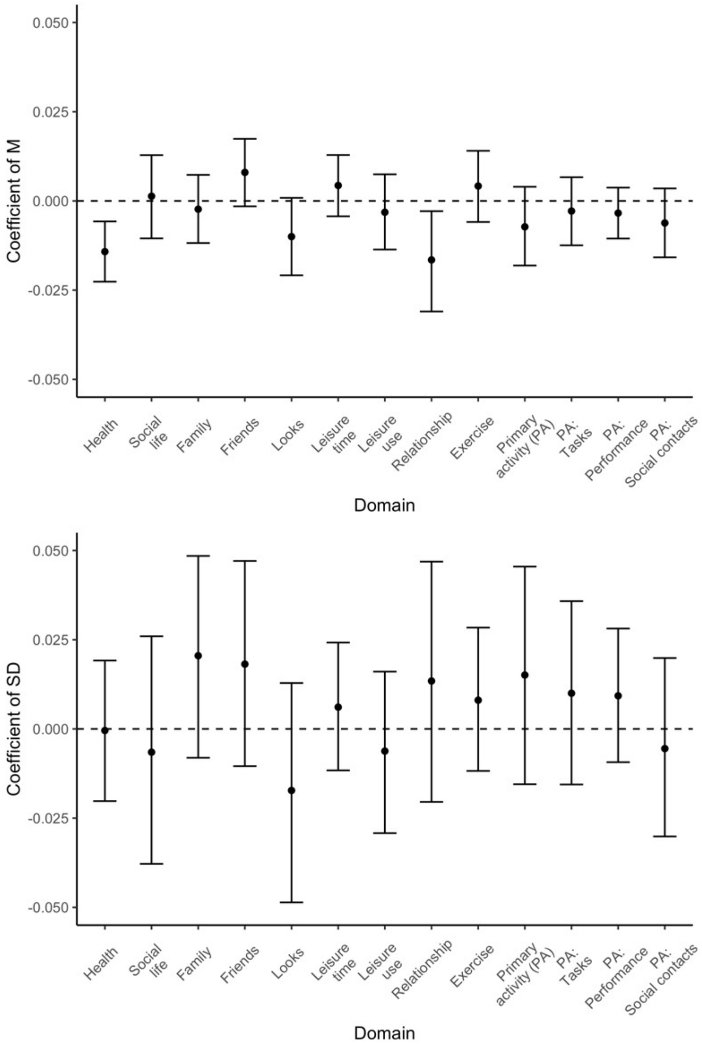

Figure 7. Associations Between Individual-Level Effects and the Person-Specific Mean and Standard Deviation of Satisfaction with the Corresponding Domain

Note. Individual-level effects were extracted from our multilevel model and subsequently regressed onto both the person-specific mean (*M*) and standard deviation (*SD*) of satisfaction with the respective domain. Error bars are 95% credible intervals.

"residual category"—if participants interpreted it to mean "anything that happened in your life which isn't work/school/studies," then any omitted life domain may induce a positive non-causal association biasing the estimated effect of leisure time use. Omitted life domains may have also biased the estimates for all other domains. While the range of domains we included is fairly comprehensive in contrast to previous studies, we did not, for example, ask participants about religious aspects of their lives, nor did we ask them about their children—a domain that is plausibly subsumed under "family life," but that in hindsight may have been interesting in its own right. Such omitted domains most likely lead to an overestimation of the effects of other domains, because factors that increase satisfaction in one domain likely also increase satisfaction in other domains, and because we can assume that the effects of domain satisfaction are positive.

Interindividual Differences in the Effects and their Correlates

Our results also suggest that there are fairly substantial interindividual differences in the effects of different life domains on general satisfaction (RQ 1). Again, we consider the overall patterns here plausible. While there was in general more interindividual effect variability for domains with larger average effects, two domains that we included precisely because we expected that they would matter for only some people (looks, exercise) showed substantial variability despite small average effects. Despite the variability in effects, individual-level effects were mostly estimated to be positive which increases our confidence in the results; a priori, it does not seem plausible that being more satisfied with any domain of life would lead anybody to be *less* satisfied with their life in general.

When it comes to correlations between interindividual differences in the effects and other variables, results get murkier. The associations that did emerge may seem sub-

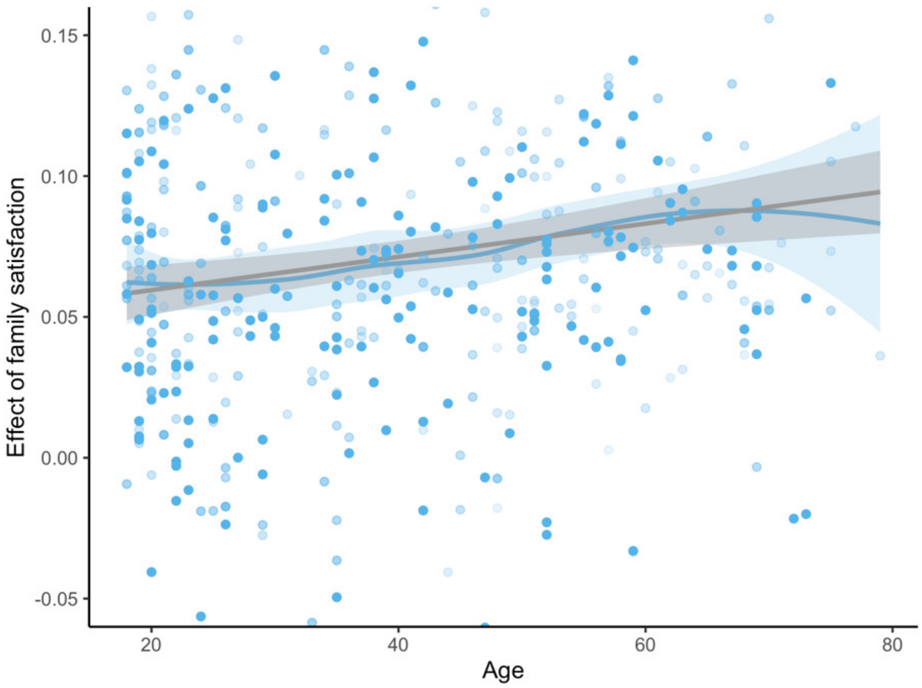

Figure 8. Association Between Participants' Age and the Estimated Individual-Level Effects of Family Satisfaction, Linear Model and LOESS.

Note. The individual-level effects of family satisfaction were extracted from our multilevel model and subsequently regressed onto age in years. Results from the linear model are presented in gray; results from locally estimated scatterplot smoothing (LOESS, default parameters) are presented in blue. Transparency of the individual data points reflects the associated estimation uncertainty; more transparent points have less influence on the analysis. Shaded area corresponds to the 95% credible interval.

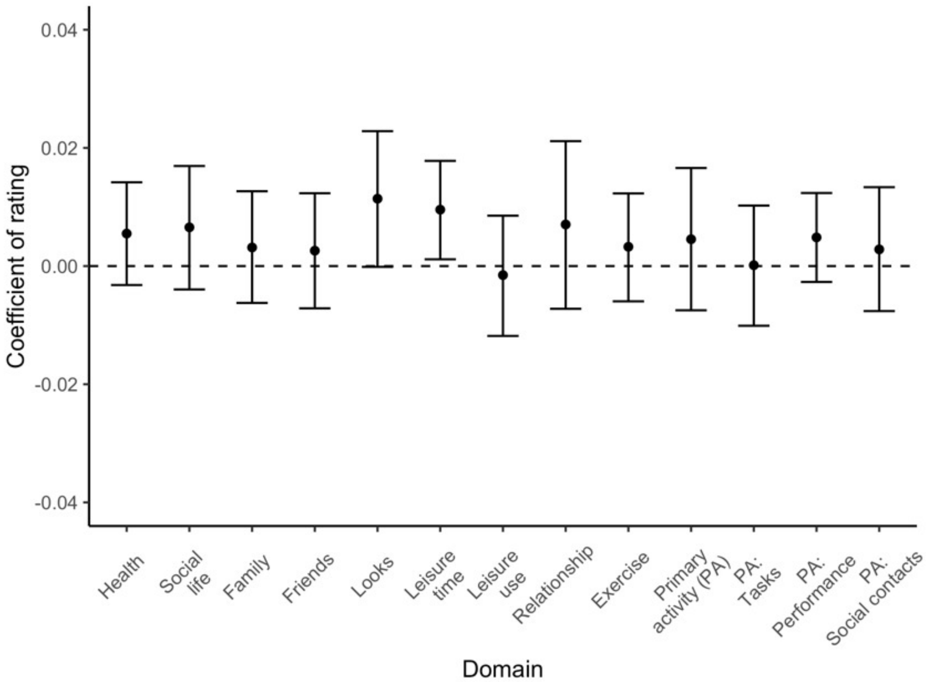

Figure 9. Associations Between Individual-Level Effects and the Combined Rating (Importance and Perceived Influence) of the Respective Domain

Note. Individual-level effects were extracted from our multilevel model and subsequently regressed onto the combined rating of the respective domain. Error bars are 95% credible intervals.

stantively plausible. For example, we found larger effects of health satisfaction for people with lower average health satisfaction and larger effects of relationship satisfaction for people with lower average relationship satisfaction—for

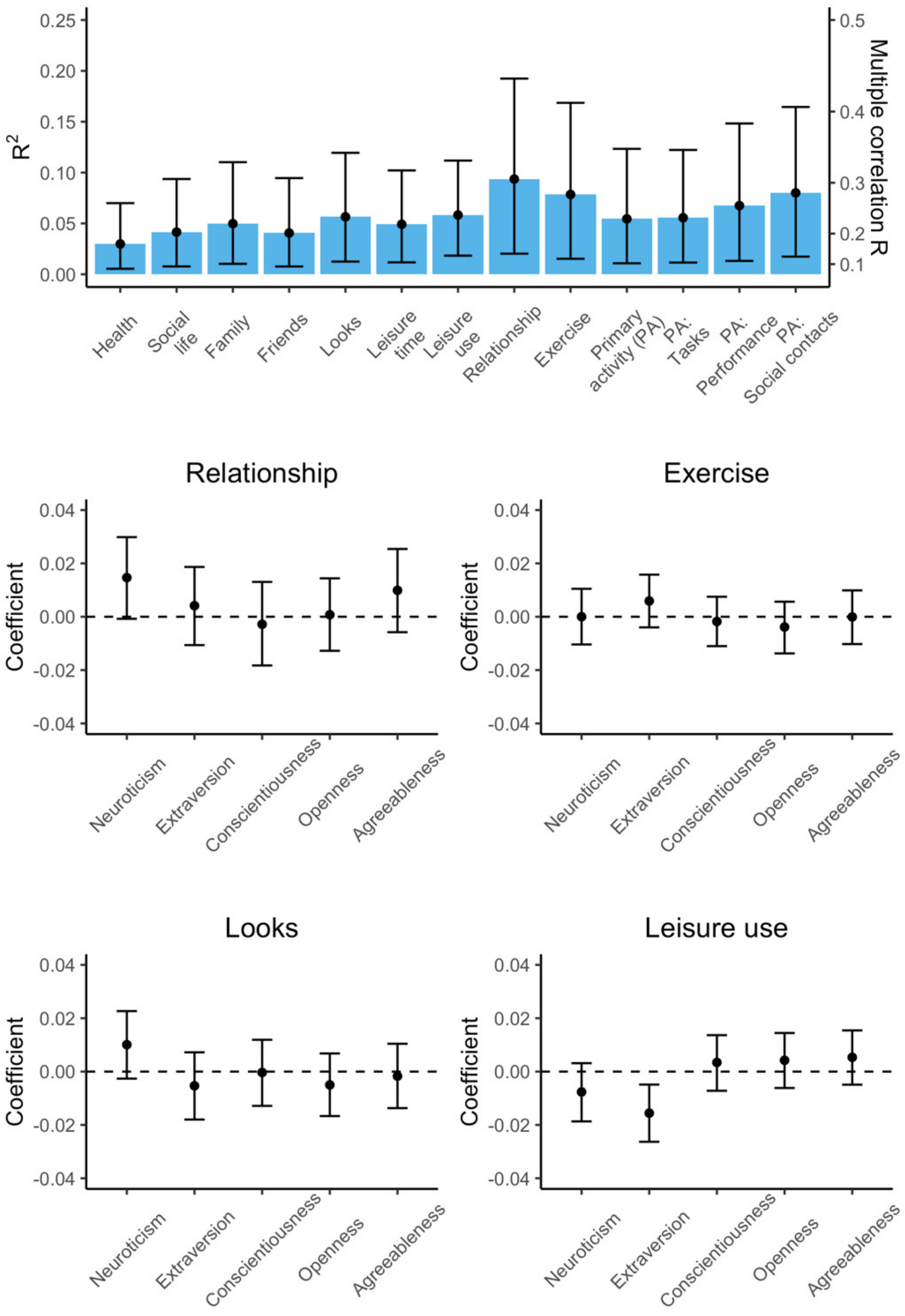

Figure 10. Associations Between Individual-Level Effects and the Big Five Personality Traits

Note. Individual-level effects were extracted from our multilevel model and subsequently regressed onto all Big Five personality traits. The upper panel presents the amount of variability in the effects of domain satisfaction on general satisfaction that the Big Five explain in combination; the panels below illustrate the model coefficients from selected domains. Error bars are 95% credible intervals.

both domains we consider it plausible that larger effects occur when things are *not* going well; follow-up analyses could investigate whether this also holds on an intraindividual level by modeling non-linear individual effects to probe whether negative deviations from one's average domain satisfaction have larger effects than positive deviations (in terms of prospect theory, whether the response to losses is stronger than the response to corresponding gains, Kahneman & Tversky, 1979). The effects of family life satisfaction were higher among older participants, which again seems plausible given that in the age range most relevant to our study, with increasing age, there is an overall shift from one's family of origin to one's own (young) children who may be much more influential for everyday life—another reason why it would have been ideal if we had asked participants about their children. We even found that among more neurotic individuals, the effects of relationship satisfaction are larger. This matches the hypothesis by Heller et al. (2006) that more neurotic individuals would be more sensitive to changes in marital satisfaction. At the same time, Heller et al. predicted the same pattern for job satisfaction, for which we found no evidence.

More generally, we ran a vast number of exploratory analyses and mostly found very little evidence for associations between the effects of domain satisfaction and other variables. But due to statistical uncertainty, we often cannot confidently rule out the existence of substantively plausible associations. For example, considering the idea of importance weighting, while across domains we found little solid evidence in favor of associations between the effects of domain satisfaction and participants' importance ratings, the estimates are still compatible with the idea that such associations exist, with positive point estimates for almost all (11 out of 13) domains. This leads us to the central limitation of our study, low statistical precision—or, in Frequentist terms, low statistical power.

Limitations

Low Power/Precision

At the time we planned the study, collecting 15 data points per participant seemed sensible to us, but this resulted in unreliable estimates of the individual-level effects. More data points would have been strongly preferable. Given that we reached reliabilities of around .30, and thinking of the diary entries as individual items, we can use the Spearman-Brown formula to predict that roughly four times as many entries would have been necessary to reach more acceptable reliabilities:

$$\begin{aligned}&\textit{Predicted reliability of an estimate based on } 60 \textit{ diary entries}\\&= \frac{4 * .30}{1 + ((4-1) * .30)} = .63.\end{aligned}$$

This happens to coincide with recommendations for Group Iterative Multiple Model Estimation (GIMME), a complex method for uncovering person-specific effects in intensive longitudinal data. The authors of GIMME suggest at least 60 time points for "adequate results" (Gates, 2023).

But more data points per participant are not the only way forward—if the goal is not the precise estimation of individual-level effects of domain satisfaction per se, but rather the precise estimation of associations between such individual-level effects and other interindividual differences, a larger number of participants can compensate for high measurement error in the individual-level effects. Ideally, one would of course increase both the number of participants and the number of data points per participant; assuming limited resources, this is probably most achievable with brief, engaging questionnaires, and a potentially open-ended study design that allows participants to provide more data points if they enjoy participating. In our study, we provided individualized feedback on well-being and the Big Five personality traits at the very end of the study; it may be worth considering intermediate feedback at multiple time points given that it is a low-cost incentive researchers can provide.

Lastly, switching from an observational to an experimental approach may improve chances to reliably estimate and potentially explain interindividual differences in individual-level effects. Our causal identification strategy necessitated the inclusion of multiple correlated predictors which increases the uncertainty in all estimates. Instead, one may provide short randomized interventions in the context of a longitudinal study (in the spirit of a "micro-randomized trial," Klasnja et al., 2015) to, for example, determine whether the effects of encouraging people to spend time with their friends vary between individuals. Such studies would target different estimands—the effects of the implemented interventions, rather than the effects of changes in domain satisfaction—which would nonetheless potentially inform us about interindividual differences in the role of life domains.

Potential Threats to Causal Identification

We already discussed the question of omitted life domains, but other types of confounding may also impair the interpretation of the effects we estimated. In the well-being literature, there is a prominent discussion about the extent to which life satisfaction judgments result from bottom-up processes (such as reflection of life circumstances, which is the premise of our models) or top-down processes (such as general personality tendencies, but also mood). While our design rules out time-invariant top-down influences (such as a general tendency to see the world positively), time-varying top-down influences (such as mood) could still introduce confounding. If such confounding by mood worked equally across domains (e.g., a given change in mood results in the same gain in satisfaction with all domains), then for each individual, we could still make statements about the differences in the effects across domains (e.g., whether for a person health satisfaction has a larger effect than family satisfaction), without being able to make statements about the absolute level of those effects. At the same time, empirical evidence suggests that individuals differ in the extent to which current emotions affect their life satisfaction (Willroth et al., 2020); thus, interindividual differences in the effects for a given domain could at least partially be attributed to interindividual differences in the amount of confounding induced by emotions. Additionally, the underlying causal structure may vary between individuals. For example, it is possible that for some people, general satisfaction causally affects domain satisfaction rather than the other way around (Beck et al., 2023). In our models, such effects would be mistaken for effects of domain satisfaction; thus, interindividual differences in causal structure may mistakenly be interpreted as interindividual differences in the magnitude of the effects of domain satisfaction.

Considerations of Time

We started from the consideration that longitudinal data help relax the necessary assumptions for the estimation of causal effects, but it is important to note that the specifics of the design and the analysis determine which effects are being estimated in the first place (Hopwood et al., 2022). First, our design could be classified as an "intensive" longitudinal design with measurement occasions spaced days apart rather than years apart. This precludes the investigation of effects that unfold over longer time scales (say, an effect of improved family life satisfaction on general life satisfaction next year), for which a traditional annual panel design would be more suited. Second, we estimated concurrent effects—effects of domain satisfaction on general life satisfaction at the very same measurement occasion—rather than lagged effects (e.g., effects of your family life satisfaction three days ago on general satisfaction now). Both of these choices are motivated from within the literature on importance weighting, in which the general idea is that respondents combine their domain satisfaction in a weighted manner when generating their general satisfaction assessment. Third, we asked participants to report on domain and general life satisfaction over the last three days (i.e., usually since the last measurement occasion). Here, we assumed that this would prompt participants to be more sensitive to changes between measurement occasions (e.g., they may be more likely to notice changes in their family lives when aggregating events across the past three days). In combination, these factors mean that our findings should not be directly generalized to make statements about potential long-term effects unfolding over longer time scales, and the extent to which they can be replicated when prompting participants only about their current satisfaction remains an open empirical question.

Conclusion

All things considered, our studies show that there is promise in the notion that different things matter to different people: we *do* find differences in the individual-level effects, and these are not just measurement error or chance fluctuations. But our findings may also temper enthusiasm to some degree.

First of all, identifying interindividual differences in effects on well-being requires a causal identification strategy that rests on additional assumptions; such strategies and assumptions are not routinely spelled out in psychology (Grosz et al., 2020). We hope that our study provides an example of how a more transparent approach to causal inference can be implemented in practice. Second, reliably estimating interindividual differences in effects does seem to require a lot of data—at least in the context of our research design. To get a better understanding of how much longitudinal data is needed to reliably estimate individual-level effects in different contexts, it would be helpful if researchers using such data routinely also reported the uncertainty associated with such individual-level effects, rather than just immediately jumping to potential explanations.

Author Contributions

Substantial contributions to conception and design: JMR ISS SCS

Acquisition of data: JRM ISS

Analysis and interpretation of data: JMR ISS RCA JS SS

Drafting the article or revising it critically for important intellectual content: JMR ISS RCA JS SCS

Final approval of the version to be published: JMR ISS RCA JS SCS

Competing Interests

We declare that there are no competing interests.

Funding

The publication of this article was supported by the Open Access Publishing Fund of Leipzig University.

Data Accessibility Statement

The complete documentation of our study including the original questions, all data and analysis scripts can be found on this paper's project page on the OSF: https://osf.io/x8j7r/.

Submitted: December 03, 2023 PDT, Accepted: June 20, 2024 PDT

References

Albers, C. J., Kiers, H. A. L., & van Ravenzwaaij, D. (2018). Credible Confidence: A pragmatic view on the frequentist vs Bayesian debate. *Collabra: Psychology*, *4*(1). https://doi.org/10.1525/collabra.149/

Arslan, R. C., Walther, M. P., & Tata, C. S. (2020). formr: A study framework allowing for automated feedback generation and complex longitudinal experience-sampling studies using R. *Behavior Research Methods*, *52*(1), 376–387. https://doi.org/10.3758/s13428-019-01236-y

Beck, E., Cheung, F., Thapa, S., & Jackson, J. (2023). *Towards personalised happiness: An idiographic approach to change in satisfaction*. https://doi.org/10.21203/rs.3.rs-3074181/v1

Böger, A., & Huxhold, O. (2018). Do the antecedents and consequences of loneliness change from middle adulthood into old age? *Developmental Psychology*, *54*(1), 181–197. https://doi.org/10.1037/dev0000453

Bürkner, P.-C. (2017). brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software*, *80*(1), 1–28. https://doi.org/10.18637/jss.v080.i01

Campbell, A., Converse, P. E., & Rodgers, W. L. (1976). *The Quality of American Life: Perceptions, Evaluations, and Satisfactions*. Russell Sage Foundation. https://play.google.com/store/books/details?id=h_QWAwAAQBAJ

Cheung, F., & Lucas, R. E. (2015). When does money matter most? Examining the association between income and life satisfaction over the life course. *Psychology and Aging*, *30*(1), 120–135. https://doi.org/10.1037/a0038682

Danner, D., Rammstedt, B., Bluemke, M., Treiber, L., Berres, S., Soto, C., & John, O. (2016). *Die deutsche Version des Big Five Inventory 2 (BFI-2)* (p. 20). ZIS - GESIS Leibniz Institute for the Social Sciences. https://doi.org/10.6102/ZIS247

Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The Satisfaction With Life Scale. *Journal of Personality Assessment*, *49*(1), 71–75. https://doi.org/10.1207/s15327752jpa4901_13

Diener, E., Lucas, R. E., & Oishi, S. (2018). Advances and Open Questions in the Science of Subjective Well-Being. *Collabra. Psychology*, *4*(1). https://doi.org/10.1525/collabra.115

Dowle, M., & Srinivasan, A. (2021). *data.table: Extension of `data.frame`*. https://CRAN.R-project.org/package=data.table

Gates, K. M. (2023, August 29). *Group Iterative Multiple Model Estimation (GIMME)*. https://cran.r-project.org/web/packages/gimme/vignettes/gimme_vignette.html

Gerson, J., Plagnol, A. C., & Corr, P. J. (2016). Subjective well-being and social media use: Do personality traits moderate the impact of social comparison on Facebook? *Computers in Human Behavior*, *63*, 813–822. https://doi.org/10.1016/j.chb.2016.06.023

Grosz, M. P., Rohrer, J. M., & Thoemmes, F. (2020). The Taboo Against Explicit Causal Inference in Nonexperimental Psychology. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, *15*(5), 1243–1255. https://doi.org/10.1177/1745691620921521

Heller, D., Watson, D., & Ilies, R. (2006). The dynamic process of life satisfaction. *Journal of Personality*, *74*(5), 1421–1450. https://doi.org/10.1111/j.1467-6494.2006.00415.x

Hopwood, C. J., Bleidorn, W., & Wright, A. G. C. (2022). Connecting Theory to Methods in Longitudinal Research. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, *17*(3), 884–894. https://doi.org/10.1177/17456916211008407

Hsieh, C.-M. (2003). Counting Importance: The Case of Life Satisfaction and Relative Domain Importance. *Social Indicators Research*, *61*(2), 227–240. https://doi.org/10.1023/A:1021354132664

Hsieh, C.-M. (2012). Should We Give Up Domain Importance Weighting in QoL Measures? *Social Indicators Research*, *108*(1), 99–109. https://doi.org/10.1007/s11205-011-9868-8

Hsieh, C.-M., & Li, Q. (2020). Validity of Importance Weighting in Subjective Well-Being Research. *The Journal of Applied Social Sciences*, *14*(2), 212–222. https://doi.org/10.1177/1936724420947336

Huxhold, O., Miche, M., & Schüz, B. (2014). Benefits of having friends in older ages: differential effects of informal social activities on well-being in middle-aged and older adults. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, *69*(3), 366–375. https://doi.org/10.1093/geronb/gbt029

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica: Journal of the Econometric Society*, *47*(2), 263–291. https://doi.org/10.2307/1914185

Kinnunen, U., Vermulst, A., Gerris, J., & Mäkikangas, A. (2003). Work–family conflict and its relations to well-being: the role of personality as a moderating factor. *Personality and Individual Differences*, *35*(7), 1669–1683. https://doi.org/10.1016/S0191-8869(02)00389-6

Klasnja, P., Hekler, E. B., Shiffman, S., Boruvka, A., Almirall, D., Tewari, A., & Murphy, S. A. (2015). Microrandomized trials: An experimental design for developing just-in-time adaptive interventions. *Health Psychology: Official Journal of the Division of Health Psychology, American Psychological Association*, *34S*(0), 1220–1228. https://doi.org/10.1037/hea0000305

Montgomery, J. M., Nyhan, B., & Torres, M. (2018). How conditioning on posttreatment variables can ruin your experiment and what to do about it: Stop conditioning on posttreatment variables in experiments. *American Journal of Political Science*, *62*(3), 760–775. https://doi.org/10.1111/ajps.12357

Oishi, S., Diener, E., Suh, E., & Lucas, R. E. (1999). Value as a moderator in subjective well-being. *Journal of Personality*, *67*(1), 157–184. https://doi.org/10.1111/1467-6494.00051

Pfund, G. N., Willroth, E. C., Mroczek, D. K., & Hill, P. L. (2024). Valuing versus having: The contrary roles of valuing and having money and prestige on well-being. *Social Psychological and Personality Science*, *15*(3), 275–287. https://doi.org/10.1177/19485506231166048

Potter, S., Gerstorf, D., Schmiedek, F., Drewelies, J., Wolff, J. K., & Brose, A. (2021). Health sensitivity in the daily lives of younger and older adults: correlates and longer-term change in health. *Aging & Mental Health*, 1–9. https://doi.org/10.1080/13607863.2021.1913475

R Core Team. (2022). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Rammstedt, B., Danner, D., Soto, C. J., & John, O. P. (2020). Validation of the short and extra-short forms of the Big Five Inventory-2 (BFI-2) and their German adaptations. *European Journal of Psychological Assessment: Official Organ of the European Association of Psychological Assessment*, *36*(1), 149–161. https://doi.org/10.1027/1015-5759/a000481

Revelle, W. R. (2017). *psych: Procedures for Personality and Psychological Research*. https://www.scholars.northwestern.edu/en/publications/psych-procedures-for-personality-and-psychological-research

Rohrer, J. M. (2018). Thinking Clearly About Correlations and Causation: Graphical Causal Models for Observational Data. *Advances in Methods and Practices in Psychological Science*, *1*(1), 27–42. https://doi.org/10.1177/2515245917745629

Rohrer, J. M., & Arslan, R. C. (2021). Precise Answers to Vague Questions: Issues With Interactions. *Advances in Methods and Practices in Psychological Science*, *4*(2), 25152459211007370. https://doi.org/10.1177/25152459211007368

Rohrer, J. M., & Murayama, K. (2023). These Are Not the Effects You Are Looking for: Causality and the Within-/Between-Persons Distinction in Longitudinal Data Analysis. *Advances in Methods and Practices in Psychological Science*, *6*(1), 25152459221140840. https://doi.org/10.1177/25152459221140842

Rohrer, J. M., & Schmukle, S. C. (2018). Individual Importance Weighting of Domain Satisfaction Ratings does Not Increase Validity. *Collabra. Psychology*, *4*(1). https://doi.org/10.1525/collabra.116

RStudio Team. (2020). *RStudio: Integrated Development Environment for R*. RStudio, PBC. http://www.rstudio.com

Scheel, A. M., Tiokhin, L., Isager, P. M., & Lakens, D. (2020). Why Hypothesis Testers Should Spend Less Time Testing Hypotheses. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, 1745691620966795. https://doi.org/10.1177/1745691620966795

Schöllgen, I., Morack, J., Infurna, F. J., Ram, N., & Gerstorf, D. (2016). Health sensitivity: Age differences in the within-person coupling of individuals' physical health and well-being. *Developmental Psychology*, *52*(11), 1944–1953. https://doi.org/10.1037/dev0000171

Simonsohn, U. (2019). *[80] Interaction Effects Need Interaction Controls*. http://datacolada.org/80

Soto, C. J., & John, O. P. (2017). Short and extra-short forms of the Big Five Inventory–2: The BFI-2-S and BFI-2-XS. *Journal of Research in Personality*, *68*, 69–81. https://doi.org/10.1016/j.jrp.2017.02.004

Westfall, J., & Yarkoni, T. (2016). Statistically Controlling for Confounding Constructs Is Harder than You Think. *PloS One*, *11*(3), e0152719. https://doi.org/10.1371/journal.pone.0152719

Węziak-Białowolska, D., Białowolski, P., & Sacco, P. L. (2019). Involvement with the arts and participation in cultural events—Does personality moderate impact on well-being? Evidence from the U.K. Household Panel Survey. *Psychology of Aesthetics, Creativity, and the Arts*, *13*(3), 348–358. https://doi.org/10.1037/aca0000180

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer. https://market.android.com/details?id=book-XgFkDAAAQBAJ

Willroth, E. C., & Atherton, O. E. (2024). Best Laid Plans: A Guide to Reporting Preregistration Deviations. *Advances in Methods and Practices in Psychological Science*, *7*(1), 25152459231213800. https://doi.org/10.1177/25152459231213802

Willroth, E. C., John, O. P., Biesanz, J. C., & Mauss, I. B. (2020). Understanding short-term variability in life satisfaction: The Individual Differences in Evaluating Life Satisfaction (IDELS) model. *Journal of Personality and Social Psychology*, *119*(1), 229–248. https://doi.org/10.1037/pspp0000261

Willroth, E. C., Pfund, G. N., Mroczek, D. K., & Hill, P. L. (2024). Perceptions of a good life: Associations with culture, age, wellbeing, and health. *European Journal of Personality*, *38*(3), 405–425. https://doi.org/10.1177/08902070231186344

Wysocki, A. C., Lawson, K. M., & Rhemtulla, M. (2022). Statistical Control Requires Causal Justification. *Advances in Methods and Practices in Psychological Science*, *5*(2), 25152459221095824. https://doi.org/10.1177/25152459221095823

Xie, Y. (2022). *knitr: A General-Purpose Package for Dynamic Report Generation in R*. https://yihui.org/knitr/

Yzerbyt, V. Y., Muller, D., & Judd, C. M. (2004). Adjusting researchers' approach to adjustment: On the use of covariates when testing interactions. *Journal of Experimental Social Psychology*, *40*(3), 424–431. https://doi.org/10.1016/j.jesp.2003.10.001

Supplementary Materials

Supplemental Material

Download: https://collabra.scholasticahq.com/article/121238-the-effects-of-satisfaction-with-different-domains-of-life-on-general-life-satisfaction-vary-between-individuals-but-we-cannot-tell-you-why/attachment/236351.docx?auth_token=Zc6MAIyLCqC2sALDPJ3D

Peer Review Communication

Download: https://collabra.scholasticahq.com/article/121238-the-effects-of-satisfaction-with-different-domains-of-life-on-general-life-satisfaction-vary-between-individuals-but-we-cannot-tell-you-why/attachment/236352.docx?auth_token=Zc6MAIyLCqC2sALDPJ3D

Response Letter

Download: https://collabra.scholasticahq.com/article/121238-the-effects-of-satisfaction-with-different-domains-of-life-on-general-life-satisfaction-vary-between-individuals-but-we-cannot-tell-you-why/attachment/236353.pdf?auth_token=Zc6MAIyLCqC2sALDPJ3D