AMERICAN PSYCHOLOGICAL ASSOCIATION

Journal of Personality and Social Psychology:
Personality Processes and Individual Differences

ISSN: 0022-3514

2024, Vol. 126, No. 4, 676–693
https://doi.org/10.1037/pspp0000501

Most People's Life Satisfaction Matches Their Personality Traits: True Correlations in Multitrait, Multirater, Multisample Data

René Mõttus[1, 2], Anu Realo[2, 3, 4], Jüri Allik[2, 4], Liisi Ausmees[2], Samuel Henry[1], Robert R. McCrae[5], and Uku Vainik[2, 6]

[1] Department of Psychology, University of Edinburgh
[2] Institute of Psychology, University of Tartu
[3] Department of Psychology, University of Warwick
[4] Estonian Academy of Sciences, Tallinn, Estonia
[5] Gloucester, Massachusetts, United States
[6] Institute of Genomics, University of Tartu

Despite numerous meta-analyses, the true extent to which life satisfaction reflects personality traits has remained unclear due to overreliance on a single method to assess both and insufficient attention to construct overlaps. Using data from three samples tested in different languages (Estonian, $N = 20{,}886$; Russian, $N = 768$; English, $N = 600$), we combined self- and informant-reports to estimate personality domains' and nuances' true correlations (r_{true}) with general life satisfaction (LS) and satisfactions with eight life domains (DSs), while controlling for single-method and occasion-specific biases and random error, and avoiding direct construct overlaps. The associations replicated well across samples. The Big Five domains and nuances allowed predicting LS with accuracies up to $r_{\text{true}} \approx .80$–$.90$ in independent (sub)samples. Emotional stability, extraversion, and conscientiousness correlated $r_{\text{true}} \approx .30$–$.50$ with LS, while its correlations with openness and agreeableness were small. At the nuances level, low LS was most strongly associated with feeling misunderstood, unexcited, indecisive, envious, bored, used, unable, and unrewarded ($r_{\text{true}} \approx .40$–$.70$). Supporting LS's construct validity, DSs had similar personality correlates among themselves and with LS, and an aggregated DS correlated $r_{\text{true}} \approx .90$ with LS. LS's approximately 10-year stability was $r_{\text{true}} = .70$ and its longitudinal associations with personality traits mirrored cross-sectional ones. We conclude that without common measurement limitations, most people's life satisfaction is highly consistent with their personality traits, even across many years. So, satisfaction is usually shaped by these same relatively stable factors that shape personality traits more broadly.

Keywords: personality traits, life satisfaction, well-being, multirater

Supplemental materials: https://doi.org/10.1037/pspp0000501.supp

General life satisfaction (LS)—an evaluative assessment of the overall degree of being satisfied with one's life (Heller et al., 2004)—is among the most desirable psychological outcomes and often an end unto itself, at least in the Western world (e.g., McMahon, 2006). Historically the purview of religion and philosophy, studying LS's causes and psychological background now involves scientists from numerous fields working worldwide (Diener et al., 2018). Much of this work has focused on LS's degree of reflecting a broader range of relatively stable psychological characteristics, besides being directly influenced by short-term situational influences and more enduring life circumstances like culture, societal and economic processes, income, health, career, relationships, and how people interpret these (e.g., Diener et al., 2018; Heller et al., 2004; Jagodzinski, 2010; Luhmann et al., 2012).

Many of the psychological characteristics are summarized with the Big Five or Five-Factor model (Costa & McCrae, 1992) or

René Mõttus https://orcid.org/0000-0002-6403-5224

This work has been supported by the Estonian Research Council's Grants PSG656 and PSG759 awarded to Uku Vainik, the Estonian Research Council's Grant PRG1291 awarded to Tõnu Esko, the Estonian Research Council's Grant PRG2190 awarded to René Mõttus, and the University of Edinburgh School of Philosophy, Psychology and Language Sciences Small Grant awarded to René Mõttus. The authors thank Rafaele Gidoro, Shivani Khanna, and Roxana Hofmann for help with recruiting and testing English-speaking participants, Wendy Johnson and Ulirch Schimmack for reading an earlier draft of the article, and Umberto Noe for statistical help. Robert R. McCrae receives royalties from the NEO Inventories.

René Mõttus played a lead role in data curation, formal analysis, methodology, writing–original draft, and writing–review and editing and an equal role in project administration. Anu Realo played a supporting role in writing–review and editing. Jüri Allik played a supporting role in writing–review and editing. Liisi Ausmees played a supporting role in writing–review and editing. Samuel Henry played a supporting role in data curation and writing–review and editing. Robert R. McCrae played a supporting role in formal analysis and writing–review and editing and an equal role in methodology. Uku Vainik played a lead role in funding acquisition, a supporting role in writing–review and editing, and an equal role in project administration.

Correspondence concerning this article should be addressed to René Mõttus, Department of Psychology, University of Edinburgh, 7 George Square, EH8 9JZ Edinburgh, Scotland. Email: rene.mottus@ed.ac.uk

HEXACO (Ashton & Lee, 2020) personality domains (Bainbridge et al., 2022). In the Big Five, neuroticism tends to have the strongest ($r \approx .40$) and openness the weakest correlation ($r \approx .10$) with LS. In the HEXACO, LS correlates the strongest with extraversion ($r \approx .40$) and the weakest with emotionality, openness, and honesty-humility ($r \approx .10$). These "Big Few" domains collectively account for about 30% of LS's variance (Anglim et al., 2020; Busseri & Erb, 2023). Insomuch as the domains represent relatively—albeit far from fully—stable individual differences, LS's correlations with them are consistent with other evidence of its traitlike nature, such as moderate long-term stability (e.g., Lucas et al., 2018), similarity among genetically related people (Weiss et al., 2008) and visibility to others (e.g., Dobewall et al., 2013; Schneider & Schimmack, 2009).

Distinct but Entangled

Regardless of its empirical correlations with personality traits, LS can remain conceptually distinct from them. On the one hand, people's differences in LS could mirror their personality traits in normal circumstances that allow them to shape and evaluate their lives according to their psychological and other traits.[1] In this case, LS can appear like any other trait—relatively stable, observable by others, and partly tracking individuals' genetic differences that provide distal backgrounds for any developmental aspect (Avinun, 2020; Bouchard, 2016; Johnson, 2010; Turkheimer et al., 2014). On the other hand, at least hypothetically, it may be possible to imagine these same people living in such dreadful circumstances (e.g., in an active war zone or concentration camp) that the majority are unhappy with most aspects of their current lives, despite still differing in some personality traits that could otherwise track LS (e.g., assertiveness or self-discipline; Anglim et al., 2020). LS's empirical associations with personality traits would then be weakened, primarily due to reduced variance in LS.

When present, empirical associations may imply personality traits' involvement in LS. For example, the traits are weakly but pervasively linked with many life outcomes that may contribute to LS, as well as with people's interpretations of their circumstances (e.g., Beck & Jackson, 2022; Soto, 2019). But this does not necessarily mean that the traits are LS's directly interpretable causes. Aspects of people's differences can become indirectly entangled over time as individuals strive toward circumstances that match their traits and possibly adapt their traits to the circumstances (Caspi et al., 2005; Johnson, 2010). For example, multiple personality traits are often linked with academic, occupational, relationship, lifestyle, and other outcomes (Seeboth & Mõttus, 2018; Soto, 2019; Stewart et al., 2022). Over time, these can further contribute to other outcomes and traits, including LS and the personality traits that influenced the outcomes in the first place (Caspi et al., 2005). Such reciprocal, crisscrossing interplay among traits and outcomes can lead to correlation patterns without easily discernible one-to-one causal relationships (Avinun, 2020). If this is the case, the overall predictability of LS from personality traits—the degree to which LS typically becomes aligned with personality—might be an equally meaningful research question compared to identifying which specific traits most closely track LS.

Here, we assumed that, in normal circumstances, LS is a relatively stable trait that both people themselves and others who know them well can evaluate with some degree of accuracy. Given this, our unique multitrait, multirater design allowed us to ask exactly how strongly LS reflects numerous other traits when controlling for previously unresolved methodological issues such as single-method and occasion-specific biases, random error, and construct overlaps. In other words, with common measurement issues eliminated, can individuals' LS be accurately predicted from their personality traits, suggesting that it is usually shaped by these same factors that shape personality traits? And if so, which traits become particularly strongly linked with LS? Or is most of LS's variance unshared with personality traits, implying that it is largely shaped by social, cultural, situational, cognitive, and other factors that have little to do with personality more broadly? These are among LS research's most fundamental questions, and accurate answers to them will necessarily constrain theorizing on LS's nature and origins (Diener et al., 2018; Heller et al., 2004). Currently, however, these answers are inconclusive despite hundreds of studies and multiple meta-analyses (Anglim et al., 2020; DeNeve & Cooper, 1998; Steel et al., 2008).

Need to Move Beyond Single-Method Studies

The typically reported correlations between LS and personality traits may misrepresent their overlaps. This is because most studies have relied on self-ratings to assess both, likely overestimating their associations due to shared single-method effects like biased self-perception or characteristic response styles (Paulhus & Vazire, 2007) that can make up much of trait score variance (McCrae & Mõttus, 2019). In cross-sectional data, correlations may also be inflated due to occasion-specific short-term effects, such as mood fluctuations or recent events. Conversely, random measurement error and raters' idiosyncratic interpretations of each construct's measures can attenuate the correlations. So, observed correlations like .10 or .40 may be either inflated estimates of much weaker or even nonexistent "true" associations, or attenuated estimates of much stronger true associations. Substantial overlaps among personality traits can further distort the (univariate) correlations because it may often be the same personality variance that is linked with LS under different trait labels (Busseri & Erb, 2023).

Combining self-reports with other information sources can help better approximate the correlations' true magnitude (Schimmack, 2010). Ratings by informants like partners, friends, or relatives provide one such source (Vazire, 2006) and show at least moderate and comparable agreement with self-reports for both LS (Schneider & Schimmack, 2009) and personality traits (Connelly & Ones, 2010). Despite numerous calls for multirater designs (Anglim et al., 2020; Diener et al., 2018), they remain rare (Dobewall et al., 2013; Schimmack et al., 2004), especially in large multisample studies that are most likely to provide robust estimates.

[1] Here, unusual circumstances would be those that are unrelated to people's own traits and that impose extreme constraints on people's freedom to live and/or assess their life according to their characteristics; examples could include active war zones, extreme societal poverty or crime, or strict pandemic lockdowns. It is likely that most of our participants did not experience such unusual circumstances, although some may have experienced acute stress stemming from sources unrelated to their own characteristics (e.g., the death of a loved one). Data from Estonian and Russian speakers were collected during mild antipandemic measures that did not restrict most individuals' freedoms. Only our English-speaking participants were tested during a stricter pandemic lockdown.

Disattenuating for Invalidity

Self-reported LS's correlations with informant-reported personality traits, and the other way around, are not inflated by shared single-method or assessment occasion-specific biases. However, they are attenuated by imperfect cross-rater agreement on both constructs—for example, due to different access to trait-relevant information or each rater's idiosyncratic interpretations of personality trait and/or LS measures, occasion-specific effects, and random error. But these factors also attenuate raters' same-trait correlations, so the ratios of average (across the two directions) cross-rater, cross-trait correlations to the average of the two cross-rater, same-trait correlations approximate traits' true associations, free of single-method and occasion-specific biases and random error.

This approach exactly parallels the familiar method of disattenuating monomethod correlations for unreliability, in which two variables' (x and y) raw correlation r_{xy} is divided by the square root of the product of their reliabilities r_{xx} and r_{xyy}:

$$r_{\text{dissattenuated}} = \frac{r_{xy}}{\sqrt{r_{xx} r_{yy}}}. \tag{1}$$

This provides an estimate of the correlation that would be observed if both measures were perfectly reliable. In the present study, we divide the cross-method, cross-variable correlation by the square root of the product of the cross-method validities, such as:

$$\frac{r_{x(\text{self})y(\text{informant})}}{\sqrt{r_{x(\text{self})x(\text{informant})} r_{y(\text{self})y(\text{informant})}}}. \tag{2}$$

A second estimate of this value is given by disattenuating the complementary cross-method correlation, $r_{x(\text{informant})y(\text{self})}$, and we define true correlations, r_{true}, as the geometric mean of these two, so:

$$r_{\text{true}} = \sqrt{\frac{r_{x(\text{self})y(\text{informant})} r_{x(\text{informant})y(\text{self})}}{r_{x(\text{self})x(\text{informant})} r_{y(\text{self})y(\text{informant})}}}. \tag{3}$$

This is the correlation that would be observed if both measures were perfectly reliable and valid.

Need to Move Beyond Broad Trait Domains

Because traits are hierarchically organized, broad domains may partly misrepresent LS's relations with personality traits. Domains can be subdivided into a few dozens of narrower traits, facets, and these further into many dozens of yet narrower traits, nuances, that also demonstrate the essential properties of traits such as relative stability over time, cross-method correlations, and partially unique etiologies (McCrae, 2015; Mõttus et al., 2019). Facets and nuances often hold unique information about life outcomes and other traits (e.g., Revelle et al., 2021; Seeboth & Mõttus, 2018; Stewart et al., 2022). LS is likely no exception, attested by its different correlations with supposedly parallel domain and facet scales that combine different nuances (Anglim et al., 2020), such as those considered similar in the Big Five and HEXACO (Thielmann et al., 2022) or assessed with different Big Five questionnaires.

LS's evaluations may directly overlap with some personality facets and nuances, hence trivially inflating their domains' correlations with LS (Steel et al., 2008; Wood & Harms, 2016). For example, LS correlates most strongly with scales asking people about self-esteem, happiness, and optimism and not feeling depressed, hopeless, and inferior to others (Anglim et al., 2020). These traits—hidden behind facet labels like depression and positive emotions—could be among life quality's definitional characteristics for many people, which would be evidenced by their r_{true}s with LS items being nearly equal to or even higher than some LS items' r_{true}s among themselves (Campbell & Fiske, 1959). Net of such directly overlapping facets and/or nuances within them, LS may reflect personality traits to a lesser degree than typical estimates show (Mõttus, 2016).

But many domains' constituent traits could also have meaningful links with LS such as agreeableness' trust facet or conscientiousness' achievement-striving and self-discipline facets, or nuances within these facets (Anglim et al., 2020). For example, a sociability facet's nuance about enjoying others' company might be more strongly linked with LS than its talkativeness nuance. Moreover, LS can be linked with specific personality traits not yet covered by most Big Five and HEXACO measures. For instance, given LS's link with relative as well as absolute income (Boyce et al., 2010; Cheung & Lucas, 2016), envy may be one narrow trait tracking low LS (Rentzsch & Gross, 2015). Or, given LS's links with having strong relationships (e.g., Diener & Seligman, 2002), low LS may have a distinct association with a tendency to feel isolated/alienated/mistreated. In this case, using only domains or even their commonly assessed facets may underestimate the overall extent to which LS reflects personality traits, let alone the associations' details. A systematic description of LS's correlations with a range of personality nuances is currently lacking, but it would help to better understand LS's broader psychological background. For example, not only LS but many other desirable life outcomes tend to go with desirable levels of (nearly) all Big Five domains (Bleidorn et al., 2020), whereas high LS may correspond to a more distinctive nuance-level profile (Stewart et al., 2022). Combining self-ratings with informant-ratings to calculate r_{true}s makes nuances' and broader traits' degrees of reflecting LS directly comparable by removing artifactual differences due to narrower trait assessments' higher measurement error. This allows nuances' distinct associations with LS to emerge more clearly, should they exist.

Need to Move Beyond a Single Way to Assess Satisfaction

If LS is defined as people's satisfaction with their lives rather than with themselves, its evaluation should reflect a broad combination of satisfactions with life's specific domains, such as work, financial and residential circumstances, and relationships (Payne & Schimmack, 2020). If so, LS should not only track a range of domain satisfactions (DSs) and especially their aggregate, but the DSs and LS should also have similar correlation patterns with personality traits. Theoretically, this could show the extent that population variance in being satisfied reflects a general trait rather than many domain-specific evaluations, possibly because the same personality traits are similarly, if indirectly, linked with how people shape different aspects of their lives and evaluate these. From a methodological perspective, assessing DSs beside LS could mitigate the risk that correlations between personality traits and satisfaction are merely due to superficial overlaps in constructs or their assessments: even if unspecific LS assessments (e.g., "Am happy with my life") may be directly based on behaviors, thoughts, and feelings also asked about to assess personality traits (e.g., "Am energetic"), this could be less likely for individual DSs (e.g., "Am happy with my relationships" or "Am happy with where I live"). Therefore, we operationalize satisfaction as both general LS and a combination of eight specific DSs, estimating their r_{true}s among each other and with personality traits.

Personality traits might track with LS more strongly than with a broad combination of DSs. This may be because people assess their overall life quality based on their personal characteristics besides their life circumstances per se (Heller et al., 2004), LS's links with personality traits are inflated by construct/measurement overlaps that researchers could not avoid, and/or researchers did not consider all relevant DSs. Therefore, we are skeptical that any given research design could fully disentangle the so-called "bottom-up" and "top-down" causal explanations (Heller et al., 2004; Payne & Schimmack, 2020) whereby, respectively, personality traits are linked with satisfaction via shaping different life domains and evaluations of these (personality traits → DSs → LS) versus primarily tracking general satisfaction that then influences satisfactions with different life domains (personality traits → LS → DSs). Besides, these explanations are not mutually exclusive (Heller et al., 2004). So, here we assessed both LS and DSs to study satisfaction's construct validity and the robustness of its links with personality traits to different ways of operationalizing it.

Need for Multisample Studies

LS's correlations with social and economic factors can vary across cultural and societal circumstances (Oishi et al., 1999; Suh et al., 1998), and so could its associations with personality traits. For example, although the domains of positive and negative emotionality, respectively resembling the extraversion and neuroticism domains, are linked with LS, these associations' strengths can vary, with the former being stronger in individualist countries and the latter in countries valuing self-expression over survival (Kööts-Ausmees et al., 2013; Kuppens et al., 2008).

So, research estimating LS's (true) associations with personality traits should also examine the findings' robustness across samples with diverse backgrounds. It is possible, for example, that narrower traits' links with LS are less generalizable than those of broad personality domains because subtle cultural and societal effects may be diluted in the domains' assessments. Likewise, using a multirater design to control for methodological issues may either dampen or magnify cross-sample variations if single-method associations have been differentially biased in different samples (e.g., random error or socially desirable responding may vary with samples). In any case, the degree of the links' robustness across samples speaks to the extent to which LS's variance reflects personality traits, besides being directly sensitive to circumstances that vary between samples and do not influence personality traits more broadly. Here, we examine the robustness of LS-personality trait links across three samples. While all samples are predominantly of European heritage, they differed in historical–societal backgrounds (e.g., historical welfare levels, political regimes or cultural influences) and languages spoken: an Estonian-speaking majority sample of Estonian residents, a Russian-speaking minority sample of Estonian residents, and a mixed-background sample of mostly Western Europeans who were tested in English.

This Study

In this largest yet multitrait, multirater, multisample study, we estimated LS's and DSs' true associations with each other and a range of broad and narrow personality traits, controlling for single-method biases, occasion-specific effects, and random error. We also avoided direct construct/measurement overlaps between personality traits and LS/DSs by ensuring that LS's indicators had higher convergent validity among themselves than discriminant validity with personality trait indicators. We additionally tested LS's true rank-order stability across several years and compared its cross-sectional r_{true}s to longitudinal ones. Specifically, 20,886 Estonian adults provided self-reports and were rated by an informant using a diverse pool of 198 items. These items were selected to cover LS and encompass a broader-than-usual range of personality traits, including the Big Five. Participants also rated their satisfaction with eight life domains: job, career choice, financial situation, residence, country, relationships, health, and appearance. In a subsample of 514 participants, personality traits and LS had also been rated by participants and their informants approximately 10 years earlier. We tested the findings' robustness among Russian speakers living in Estonia ($N = 768$) and English-speaking participants from various mostly European countries ($N = 600$). All this allowed us to estimate satisfaction's overall extent of reflecting personality traits and the associations' details with a level of precision and robustness rarely, if ever, attained yet.

We may already know that some personality traits' correlations with LS are greater than zero, at least in usual circumstances. However, a more important but not yet compellingly answered question is: how much greater? For example, it would be a twofold difference if LS could be predicted from personality traits with an accuracy of .80–.90 rather than an accuracy of .50–.60,[2] and we should care about such a difference just as much as natural scientists would care whether the speed of light is 1.5×10^8 or 3×10^8 m/s or whether the Earth's atmosphere contains about 21% or 11% of oxygen. Thus, while our empirical work is descriptive and predictive (Mõttus et al., 2020), the findings significantly contribute to our theoretical understanding of satisfaction's relatively stable psychological basis.

Method

Transparency and Openness

Our sample sizes were determined by practical constraints rather than power calculations, but collectively provide high power for any nontrivial effect sizes. We report all data exclusion criteria and variable manipulations. We make our data analytic (R) scripts publicly available, as well as data from one (English-speaking) sample (Mõttus, 2023). Other data cannot be made publicly available due to being part of a large and ongoing biobank study, but researchers can apply for access at https://genomics.ut.ee/en/content/estonian-biobank. Data used in the Supplemental Analyses are also publicly available. All statistical analyses were carried out with R language, Version 4.3.1 (R Core Team, 2023). The analyses were not preregistered. Supplemental material, supplemental analyses, and supplemental tables can be found at the journal's website and in Mõttus (2023).

[2] Correlations have a nonlinear scale. To make them comparable, they have to be z-transformed.

Participants

The Estonian- and Russian-speaking data collections were approved by the Estonian Committee on Bioethics and Human Research. The Estonian and Russian speakers were members ("gene donors") of the Estonian Biobank, a population sample of approximately 200,000 adults encompassing about 20% of Estonian adult residents or past residents currently living abroad (https://genomics.ut.ee/en/content/estonian-biobank). Data used for this study were collected through an online survey between November 2021 and April 2022 and participants could choose to participate in either Estonian or Russian, most likely depending on their native language (Vaht et al., 2024). Because Estonia has a substantial Russian-speaking minority with a somewhat distinct cultural and historical background, we treated Estonian and Russian speakers as separate samples. For example, although most Russian speakers were likely born or had been living in Estonia for many years and were integrated with the Estonian society, many Russian speakers are geographically concentrated, follow different (often Russian) media and have partly distinct identities (Vihalemm et al., 2019); these are also among likely reasons that Russian speakers are underrepresented among the gene donors. Email invitations were sent to 182,405 gene donors, with up to two follow-up invitations as necessary. Participants who completed the survey were offered feedback on their Big Five personality trait scores. To encourage participation, the study was promoted on national radio, television, newspapers and magazines, and on social media. Participants were optionally asked to provide an email of another person (informant) who could complete the third-person form of the personality item pool about them. After reading information about the study, both participants and their informants electronically signed a consent form.

Estonian Speakers

In total, $N = 73{,}266$ Estonian-speaking participants completed the survey. After removing participants who either did not invite an informant or whose informant did not submit their responses, and participants with more than 10 missing responses in either self- or informant-report surveys, we were left with 20,886 participants (sex assigned at birth: 14,228 women, 6,658 men; age: range from 18 to 93; $M = 44.0$, $Mdn = 45.2$, $SD = 13.7$). The included and excluded participants somewhat differed in their average personality traits and LS (Mõttus, 2023); the 52,380 excluded participants were less open and life-satisfied than their 20,886 included peers (respectively, $d = -0.25$ and -0.14, $p < .001$), while differences in their other traits were negligible ($0.01 \leq |d| \leq 0.06$). The informants were usually partners or spouses, children/grandchildren, friends, or parents/grandparents (56%, 14%, 14%, and 7%, respectively). Between 2008 and 2017, 514 of the participants (321 females; age: range from 18 to 79 years at the time; $M = 38.7$, $Mdn = 38.0$, $SD = 13.3$) had completed another personality test and answered to an LS question. As 79% of them had participated by 2012, mostly from 2009 to 2010, and further 15% participated in 2013, the typical retesting interval was about 10 years.

Russian Speakers

Of the 3,719 Russian-speaking participants who completed the survey, we could retain data for 768 after applying inclusion criteria identical to those applied to Estonian speakers (sex assigned at birth: 533 women, 235 men; age: range from 18 to 88; $M = 43.4$, $Mdn = 43.0$, $SD = 13.0$). Akin to Estonian speakers, Russian-speakers' informants were typically partners or spouses, children/grandchildren, friends, or parents/grandparents (54%, 15%, 16%, and 8%, respectively).

English Speakers

Between March and June 2020, 300 dyads completed personality and LS items about themselves and the other dyad member in English (436 females, seven preferred not to say; age: range 12–82 years; $M = 28.5$, $Mdn = 23.0$, $SD = 12.9$). These data were originally collected for student projects exploring items' cross-rater correlations, approved by the University of Edinburgh institutional review board. People were recruited online, initially through the students' and a volunteer research assistant's personal networks, and later through announcements on various social network sites. Participants were offered feedback on their Big Five traits and most salient personality nuances, and how well they and their informant agreed regarding each other's traits. The participants recruited through social media announcements were also offered gift cards or PayPal transfers (£5). The person who started the study participation was asked to identify another dyad member and provide their email, who was then invited to similarly participate. Although the study was intended for adults, six participants invited adolescent dyad members. Most participants were British residents (57%), but many resided in other European countries (e.g., Italy: 7%; France: 3%; Spain: 3%; Greece: 2%; Poland: 2%), United States (5%), or India (5%).

Measures

Main Data Collection (2021–2022)

The 100 Nuances of Personality (100-NP), completed by people themselves and their informants, is a 198-item pool designed to cover personality traits and LS reliably, comprehensively and with reduced redundancy. It captures trait content associated with most facets and domains assessed in standard Big Five measures as well as some individual differences measures beyond these (e.g., LS, competition, envy, humor, sexuality, spirituality, and the "Dark Triad" traits). The 100-NP items were iteratively selected from larger item pools such as the International Personality Item Pool (Goldberg, 1999) and Synthetic Aperture Personality Assessment (Condon & Revelle, 2016) for their content, and retained if they (a) had acceptable levels of empirical properties (e.g., test–retest reliability, variance, and cross-rater agreement) and (b) were not redundant with other items. However, we included some highly similar items to allow for assessing acquiescent responding and r_{true}s, or to provide a pair of items for apparently less reliably assessable traits, such as impulsiveness. Besides completing the 100-NP, Estonian- and Russian-speaking participants completed five items about satisfaction with job, choice of career, financial situation, residence, and country; these items were only completed by participants themselves due to the limited number of items that could be administered to informants. All items were responded to using a 6-point Likert scale from *completely inaccurate* to *completely accurate*. Missing responses were replaced with the median. A detailed description of the 100-NP's development can be found in Henry and Mõttus (2022) and the full item list is in the Supplemental Material. Retaining items for LS, DS, and personality trait assessment is described after introducing necessary statistical analyses.

Earlier Data Collection (2008–2017)

Personality traits were measured with the Estonian version of the NEO Personality Inventory–3 (NEO-PI-3; McCrae et al., 2005). The NEO-PI-3 items were responded to using a 5-point scale, and domains and facets were scored as sum scores of their items, as per test manual. LS was assessed with a single item: "All things considered, how happy are you with your life generally?," rated on a 10-point scale from *not at all* to *completely*.

Analyses to Estimate True Associations

True Correlations (r_{true}s)

To estimate variables', say x and y, r_{true}s, we correlated self-reported x with informant-reported y and vice versa, and calculated the geometric mean of these cross-rater, cross-variable correlations. We then correlated self-reported x with informant-reported x and the same for y, and subsequently calculated the geometric mean of these cross-rater, same-variable correlations. We treated the ratio of the former geometric mean to the latter as the r_{true} between x and y, free of single-method biases that are either specific to either x or y or shared among them, rating occasion-specific effects (e.g., mood) and random error because these four variance components would similarly affect both cross-rater, cross-item and cross-rater, same-item correlations and therefore cancel out in their ratio. The approach is based on the simplifying assumption that both variables' valid (true) variance is at least partly shared between raters—hence, partly independent of assessment method—whatever its fraction to total variance, and that rating biases and occasion-specific effects are not shared between raters. The degrees of rater- and occasion-specific effects and random error may differ across variables and raters, but as long as all four correlations are used, they are equally represented in both the numerator and denominator of the r_{true} calculation and hence cancel out in equal proportions. Among other things, this means that raters' asymmetrical information about the traits does not influence the model's estimates. An extended, algebraic and graphical formalization of the variance decomposition model underlying the r_{true} calculation is in Supplemental Material. The idea is similar to how Wood et al. (2023) used test–retest data to estimate items' semantic similarity, except that we used informant-ratings instead of retest scores, which allowed us to additionally control for single-method effects.

True Predictive Accuracy

To estimate personality traits' true combined overlap with LS (unbiased "multiple R") in the Estonian-speaking sample, we created elastic net models tailored to maximize the traits' out-of-sample predictive accuracy for aggregate LS in one sample partition (67%) and calculated the correlation between LS and its values predicted from personality traits using this model in another sample partition (33%). The elastic net models with .50 α parameter were trained to minimize prediction error in 10-fold cross-validation within training samples. For true predictive accuracy, net of single-method biases and random error, we "cross-predicted" self-reported LS from informant-reported personality traits and vice versa in 10 random training-validation sample splits and averaged the predictive accuracies within each direction, and divided the geometric means (across directions) of these cross-prediction accuracies by the geometric means of self-informant correlations for (a) observed LS scores and (b) their predicted-from-personality values. For replications in Russian- and English-speaking samples, we used models trained in the Estonian data, hence training and validating the same models across different languages.

Domain Satisfactions

Because most DSs were assessed with only self-reports, we approximated their r_{true}s with LS by, first, correlating self-reported DS items with the informant-reported LS aggregate and then dividing these correlations by the geometric mean of (a) average cross-rater correlation of three DS items for which cross-rater data were available (as a proxy for all DS items' cross-rater correlation; .44) and (b) the LS aggregates' cross-rater correlation. Likewise, we approximated the DS aggregate's and LS aggregate's r_{true} by calculating self-reported DS aggregate's correlation with informant-reported LS aggregate and dividing this by the geometric mean of the LS aggregate's cross-rater correlation and the cross-rater correlation of the principal components of the three DS items for which cross-rater data were available. Because not all cross-rater correlations were used in these calculations, and ratings of different variables and/or by different raters could contain somewhat different degrees of biases and errors that were then not equally represented in the numerators and denominators of the r_{true}s approximations, the r_{true} estimates pertaining to DSs could be to some degree biased, unlike the LS-personality trait r_{true}s based on four correlations each.

Standard Errors

Because most r_{true}s were based on four correlations each, the usual standard error formulas did not apply to them. To find a formula to estimate the standard errors, we relied on an iterative process of inductive reasoning and tinkering, comparing the results against the ground truth in simulated data until the formula results closely approximated the simulation results (see Supplemental Material). The main sample of Estonian speakers was so large that the standard errors were bound to be small, but they were larger for estimates in smaller Russian- and English-based samples.

Variable Selection and Aggregation

To proof-of-principle test whether the cross-informant design could approximate variables' r_{true}s, we considered some pairs of highly similar items (e.g., "Keep my promises" vs. "Break my promises"). In the main, Estonian-based data, these items' $|r_{\text{true}}\text{s}|$ reached .97, providing support for the research design (see Supplemental Table S1 for the 100 highest-correlating item pairs). After considering the items' content (i.e., face validity), we used r_{true}s for variable selection and aggregation, unless said otherwise, and relied on the general idea that items measuring the same construct should have stronger correlations than items measuring different constructs (Campbell & Fiske, 1959). We describe these analyses based on the Estonian-based data, but Supplemental Tables S2–S6 also contain correlations for Russian- and English-based analyses, as applicable.

Life Satisfaction

Relying on face validity, we initially designated four 100-NP items to capture LS because these either directly assessed being happy with life or required people to evaluate their lives' different aspects: past, current direction, and future. The $|r_{\text{true}}\text{s}|$ among three of these items ("Am happy with my life," "Feel that my life lacks direction," "Have a dark outlook on the future") varied narrowly between .74 and .77 (see Supplemental Table S2 that also includes the items' descriptive statistics). For comparison, their single-method absolute correlations varied between .48 and .52 in self- and informant-reports. The fourth item that pertained to evaluating the past, "Life has been kind to me," had lower $|r_{\text{true}}\text{s}|$ with other items (.32–.56), so we removed it from further analyses to retain LS high construct validity (its absolute single-method correlations varied from .16 to .41). The cross-rater correlations of the three retained LS items were .42, .37, and .36. Separately in self- and informant-ratings, we used scores of the first principal components of the three LS items as aggregate LS scores (respectively, explaining 66% and 68% of the items' variance; all loadings > |.80|; cross-rater correlation .48). The items' correlation pattern replicated in Russian- and English-speaking samples (Supplemental Table S2).

In Supplemental Analyses 1, we show that latent trait scores based on these three items correlated highly (r = .95, .90, and .80, respectively, among Estonian, Russian, and English speakers) with latent trait scores of a more widely used LS assessment, Satisfaction with Life Scale (SWLS; Diener et al., 1985), based on additional self-reported data collected in Estonian, Russian, and English. Although we acknowledge that our approach to LS assessment is unconventional, the very high correlations with the SWLS strongly support the concurrent validity of our aggregate LS scores—at least, it was a very close proxy measure of LS.

Domain Satisfactions

Besides the five items completed only by Estonian- and Russian-speaking participants themselves to assess their DSs, we also designated three of the 100-NP items to capture DSs about satisfaction with relationships, health, and appearance: "Am satisfied with my relationships," "Consider myself healthy for my age," and "Consider myself good-looking." Correlations among the eight self-report items selected to assess DSs (see Supplemental Table S3 that also includes the items' descriptive statistics) were lower than those for LS items, varying from .08 to .62 (*Mdn* = .20, compared to the respective *Mdn* = .50 for the three LS items; Supplemental Table S3), but were mostly within the recommended range of a typical scale's interitem correlations (Clark & Watson, 1995). So, we used the scores of the items' first principal component as an aggregate DS score (explaining 34% of items' variance; all loadings >.40). The association pattern replicated among Russian speakers (Supplemental Table S3). In Supplemental Analyses 1, we show that latent trait scores based on these eight DS items correlated extremely highly (r = .96) with latent trait scores of the SWLS, based on additional self-reported data collected among Estonian speakers; in these data, the DS and LS latent factors also correlated r = .87. This evidence clearly supports the validity of the DS aggregate as a measure of LS.

Personality Items

We dropped personality items that comparatively more strongly overlapped among themselves (51 items) or with any LS item (one item: "Tend to feel very hopeless"), using $|r_{\text{true}}| \geq .75$ as the cutoff (given the $|r_{\text{true}}| \approx .75$ among LS items) and dropping the weaker LS correlate from each pair of highly correlating items (Supplemental Table S4 for the main Estonian-based analyses and replications in Russian and English). We also dropped three items ("Worry about my health," "Worry a lot about my looks," "Wear stylish clothing") that could semantically overlap with DSs about health and appearance. We treated the remaining 136 personality items as markers of partly distinct personality nuances that could have discriminant validity for LS and/or DS. Their cross-rater correlations ranged from .15 to .64 (*Mdn* = .30). English-speaking participants were not administered six personality items, two of which were among the 136 retained items ("Work on improving myself," "Try to provoke others just to get a response"). Conceptually, we do not treat items as nuances per se, but as markers for both broad traits like personality domains and narrow traits like nuances. However, we do use item-level correlations to describe the nuancedness of LS's personality correlates, where evidence for it exists.

Personality Domains

So far, there is no universally agreed organization of nuances (or more concretely, items) into facets and domains (Condon et al., 2020). Therefore, we skipped the facet level and combined the 136 personality items into five domains by performing a principal component analysis on their $r_{\text{true}}\text{s}$ (Supplemental Table S5). After varimax rotation, we retained 15 highest loading items from each component to ensure a roughly balanced and most relevant content representation for each domain, and recalculated the five components based on the 75 remaining items; these accounted for 60% of items' variance. After varimax rotation, these components' loadings clearly resembled the typical Big Five themes (Supplemental Table S6), and we used them to separately but identically calculate domain scores in self- and informant-ratings, multiplying standardized item scores by the items' inverted correlation matrix and the principal component loadings. Had we chosen more items per component, the loading pattern would have started differing from what we considered typical Big Five content and some loadings would have dipped below .40. This procedure ensured that domain scores were calculated similarly in self- and informant-reports. The domains' cross-rater correlations were .56 (emotional stability), .58 (extraversion), .57 (openness), .46 (agreeableness), and .50 (conscientiousness), which are comparable or higher than usual (Connelly & Ones, 2010). In single-method designs using common Big Five instruments, the domain scores can correlate as highly as .40 s and .50 s (van der Linden et al., 2010). Intentionally and desirably, our domain scores' correlations were lower, varying from $|r|$ = 0 to .12 (*Mdn* = .04) in self-reports and from $|r|$ = 0 to .21 (*Mdn* = .02) in informant-reports; $|r_{\text{true}}\text{s}|$ varied from .01 to .31 (*Mdn* = .07). This relative independence of domain scores ensured that the domain-LS correlations would be less inflated by the domains' shared variance than usual. Our longitudinal data allowed us to estimate the domains' empirical similarity to those of a widely used Big Five assessment, NEO-PI-3, although the assessments were separated by a decade (r_{true}s = .73–.82; Table 1).

Results

The following five sections describe results from the largest, Estonian-speaking sample.

Correlations With Personality Items

We started with LS's correlations with the 136 items retained as possible markers of personality nuances with discriminant validity for LS and DS, calculating their r_{true}s with individual LS items and the aggregate of these. The three vectors containing the LS items' z-transformed r_{true}s with the personality items were highly similar ($r \geq .94$), supporting the LS aggregate's construct validity. For the LS aggregate, r_{true}s varied from −.69 to .60 ($|r_{\text{true}}|_{\text{median}} = .22$; $|r_{\text{true}}|_{\text{min}} = .01$). Figure 1 shows the 70 items correlating with the aggregate LS at $|r_{\text{true}}| \geq .20$ ($|r_{\text{true}}|_{\text{median}} = .33$), while all 136 r_{true}s are in Supplemental Table S7, alongside their underlying cross-variable, cross-rater correlations and same-variable, cross-rater correlations.[3,4] For comparison, single-method correlations for these 70 items with the LS aggregate (Supplemental Table S7) varied from −.45 to .46 in self-reports ($|r|_{\text{median}} = .17$) and from −.50 to .44 in informant-reports ($|r|_{\text{median}} = .20$), so $|r_{\text{true}}\text{s}|$ tended to be stronger despite not being influenced by single-method biases.

For interpretation ease, we also highlight the LS aggregates' strongest relatively unique correlates, showing the 19 items not having $|r_{\text{true}}\text{s}| > .50$ with any other personality item with a larger and darker font in Figure 1. Because these 19 items were comparatively less intercorrelated, they covered a broader range of traits than our Big Five domains (12 were not included among the 75 Big Five items). Low LS tracked with feeling misunderstood ($r_{\text{true}} = -.69$), lack of excitement (−.61), indecisiveness (−.51), envy (−.49), boredom (−.45), and feeling used (−.41), whereas high LS tracked confidence in ones' abilities (.44) and believing that effort is rewarded (.40). Less strongly ($.20 < |r_{\text{true}}| < .40$), high LS tended to be uniquely characterized by taking risks, finding it easy to apologize, feeling special commitment to one's family, being loyal, respecting authority, liking to visit new places, and working on self-improvement, whereas low LS tended to go with making enemies, telling lies, forgetting things, and crying easily.

Table 1
Longitudinal Correlations

Construct	Later LS		Earlier LS		Correlations with earlier domain	
	r_{true}	*SE*	r_{true}	*SE*	r_{true}	*SE*
Later domain (100-NP)						
Emotional stability	.47	.046	.40	.049	.73	.038
Extraversion	.43	.045	.38	.049	.74	.033
Openness	.02	.069	.07	.067	.82	.035
Agreeableness	.04	.074	.06	.075	.82	.061
Conscientiousness	.30	.054	.12	.066	.76	.047
Earlier domain (NEO-PI-3)						
Emotional stability	.53	.045	.59	.047		
Extraversion	.43	.042	.39	.045		
Openness	.13	.058	.08	.064		
Agreeableness	.03	.077	.06	.077		
Conscientiousness	.29	.053	.30	.055		
Later LS			.70	.056		

Note. For consistency, we reverse-keyed NEO-PI-3's neuroticism as emotional stability. LS = general life satisfaction; *SE* = standard error; earlier = measured from 2008 to 2017 (mostly before 2013); later = measured from 2021 to 2022; 100-NP = 100 Nuances of Personality; NEO-PI-3 = NEO Personality Inventory–3.

Correlations With Personality Domains

Next, we correlated LS with Big Five domains to represent the LS's personality correlates more parsimoniously and comparably with typical findings in the existing literature. The domains' r_{true}s with the LS aggregate (Table 2) ranged between .30 and .47 for conscientiousness, extraversion, and emotional stability, but remained below .05 for openness and agreeableness.[5] For comparison, single-method correlations of the five domains with the LS aggregate were r = .34, .36, .10, .11, and .28 in self-reports and .41, .34, .07, .09, and .27 in informant-reports, respectively for emotional stability, extraversion, openness, agreeableness, and conscientiousness (Supplemental Table S8). So, domains' r_{true}s with LS were comparable or higher than single-method correlations for emotional stability, extraversion and conscientiousness, despite not being influenced by variables' shared single-method biases. These correlations are also comparable or even higher than those reported in other single-method studies (Anglim et al., 2020; Table 2), despite our domain scores being less intercorrelated than those in studies using common Big Five scales. For openness and agreeableness, r_{true}s were lower than single-method correlations in this and other data (Supplemental Table S8), possibly because of not being inflated by single-method biases and/or overlaps with other personality domains.

However, items primarily loaded on by the same domains often varied considerably in their correlations with the LS aggregate, such as "Enjoy hurting others" ($r_{\text{true}} = -.32$) and "Believe that

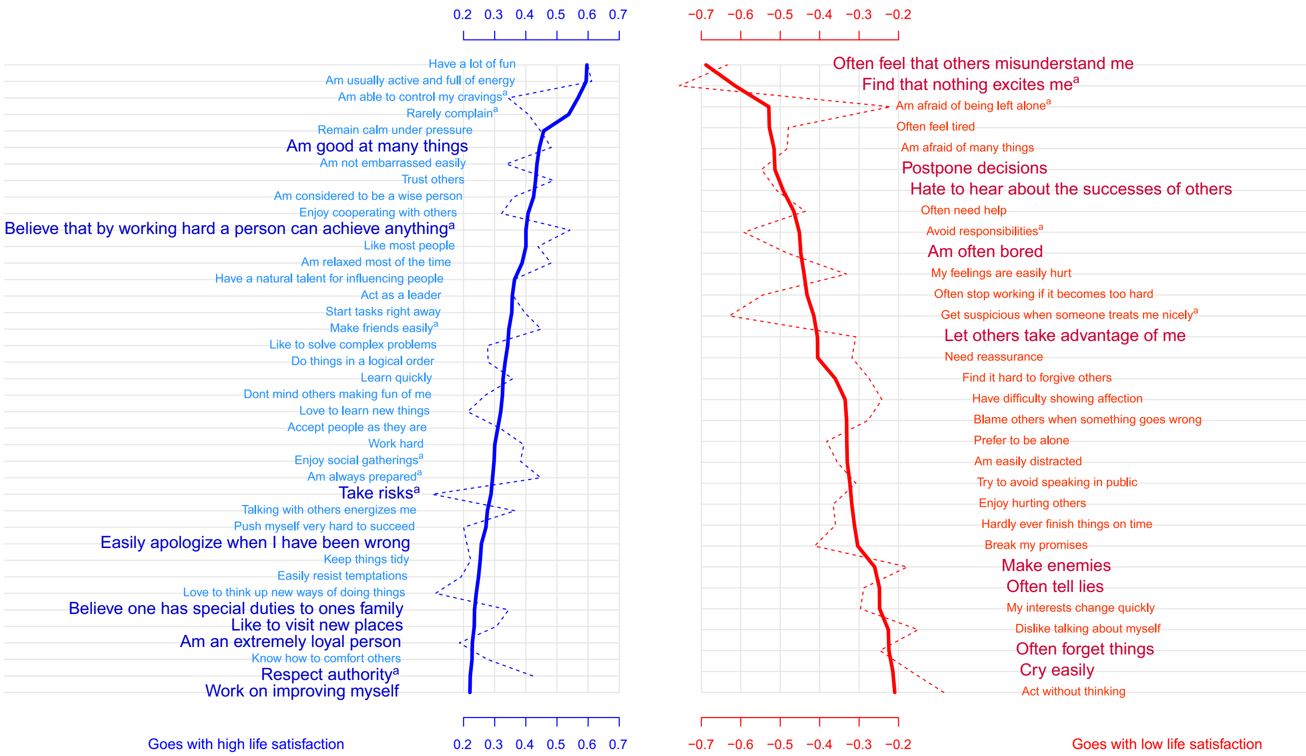
[3] The .20 cutoff was chosen for presentation ease and because correlations of .20 and higher are said to heuristically represent "a medium effect that is of some explanatory and practical use even in the short run" (Funder & Ozer, 2019, p. 156). Even if significant, small correlations among psychological variables may sometimes reflect the pervasive "crud factor," thus not being meaningfully interpretable.

[4] The two directions of cross-variable, cross-rater correlations (self-rated LS's correlations with informant-rated personality traits and the other way around; Supplemental Table S7) were highly similar, with the z-transformed correlation vectors from the two directions correlating .98/.99 for the 136/70 items. This suggests that self- and informant-rated items contain broadly similar degrees of information about the variables involved, including LS. This mitigates the possibility that LS's informant-reports are (more) biased (than self-reports).

[5] The two directions of calculating the cross-variable, cross-rater correlations (self-rated LS's correlations with informant-rated personality domains and the other way around) yielded broadly similar, although not identical results. Informant-related LS's correlations with self-rated personality domains were .28, .20, .00, .03, and .14, while self-rated LS's correlations with informant-rated domains were .21, .25, .02, −.01, and .15, respectively, for emotional stability, extraversion, openness, agreeableness, and conscientiousness.

Figure 1

Personality Items' True Correlations (r_{true}s) With the Aggregate LS in the Estonian-Speaking (N = 20,886; Solid Lines) and Combined Russian- and English-Speaking (N = 768 and 600; Dashed Lines) Samples



Note. LS = general life satisfaction. Items with greater and darker font size did not have r_{true}s higher than .50 with any other personality items among the Estonian speakers, thus reflecting relatively distinct personality nuances. See the online article for the color version of this figure.
[a] Indicates 99% confidence intervals of the combined r_{true} estimate from the English- and Russian-speaking samples did not span the estimate from the Estonian-speaking sample, suggesting potentially meaningful uniqueness in the latter. Standard errors are in Supplemental Table S7.

I am always right" (r_{true} = .03) that were both negatively loaded on by the agreeableness domain (Supplemental Table S6). This partly explains why domains $|r_{true}\text{s}|$ were lower than those of several items within and beyond the domains. So, although domains provide a parsimonious representation of LS's associations with personality traits, they can also partly misrepresent these associations.

Life Satisfaction's Overall Predictability

Next, we evaluated the overall degree to which life satisfaction aligns with individuals' personality traits. Predicting LS from the full set of 136 items and five domains, the respective true predictive accuracies were .88 and .79 (Table 3). Even though disattenuated for

Table 2

LS's True Correlations With Personality Domains in Estonian-/Russian-/English-Based Data

Domain	Estonian-based data		Russian-based data		English-based data		Meta-analytically combined (Russian and English)		Past results (for reference)
	r_{true}	*SE*	r_{true}	*SE*	r_{true}	*SE*	r_{true}	*SE*	*r* (single method)
Emotional stability	.47	.008	.36	.053	.32	.053	.34	.038	.39
Extraversion	.43	.008	.50	.052	.45	.048	.47	.035	.32
Openness	.02	.011	−.05	.066	.00	.064	−.03	.046	.08
Agreeableness	.04	.012	.02	.072	.10	.067	.06	.049	.20
Conscientiousness	.30	.010	.47	.052	.26	.060	.38	.039	.27

Note. For single-method and cross-method correlations, see Supplemental Table S8. Past results (for reference) = meta-analytic estimates from "Predicting Psychological and Subjective Well-Being From Personality: A Meta-Analysis," by J. Anglim, S. Horwood, L. D. Smillie, R. J. Marrero, and J. K. Wood, 2020, *Psychological Bulletin*, *146*(4), p. 298 (https://doi.org/10.1037/bul0000226), based on self-report data. Copyright 2020 by the American Psychological Association. LS = general life satisfaction; *SE* = standard error.

Table 3
LS's True Out-of-Sample Predictability From Personality Domains and Items in Estonian-/Russian-/English-Based Data From Models Trained in Estonian Data

Domain	Estonian-based data		Russian-based data		English-based data	
	r_{true}	*SE*	r_{true}	*SE*	r_{true}	*SE*
Five domains	.79	.008	.74	.046	.64	.049
136/134 items[a]	.88	.007	.90	.040	.84	.035
70/69 items[a]	.87	.008	.86	.050	.82	.037
19/18 items[a]	.86	.010	.88	.050	.82	.042
Three items	.81	.010	.82	.057	.82	.046

Note. r_{true} = true correlation between predicted and observed life satisfaction. LS = general life satisfaction; *SE* = standard error.
[a] Smaller item numbers apply to English-based data.

measurement issues, these represent unusually high correlations in psychological research (Funder & Ozer, 2019).[6] To see whether the estimates were driven by numerous predictors—either by many items individually or by many items contributing toward domain scores—we also explored true predictive accuracies of smaller item sets such as the 70 items shown in Figure 1, the 19 relatively unique items among them (not having $|r_{true}s| > .50$ with any other items), and the three most strongly LS-related items among these 19 ("Often feel that others misunderstand me," "Find that nothing excites me," and "Postpone decision"). These smaller item subsets provided true predictive accuracies between .81 and .87 (Table 3), showing that LS was highly predictable from even a few specific personality traits. For comparison, the predicted-observed LS correlations in single-method data were .75 and .64 in self-reports and .76 and .64, in informant-reports, respectively, for models based on 136 items and five domains. For domains, this corresponds to $R^2 = .41$, which is higher than the $R^2 \approx .30$ usually found in single-method studies (Anglim et al., 2020; Busseri & Erb, 2023).

Longitudinal Analyses

We used the longitudinal assessments in the Estonian-speaking sample to assess the stability of LS and its personality correlates over time. Should LS's cross-sectional and longitudinal r_{true}s with personality traits be similar and approach LS true stability, this would suggest that personality traits' systematic involvement in LS endures over time, irrespective of time-varying influences on either.

The two Big Five domains' assessments, separated by approximately 10 years, had r_{true}s between .73 and .82, and the single-item LS had the r_{true} of .70 with LS 10 years later (Table 1).[7] The later Big Five scores correlated with the earlier and later LS similarly for all domains but conscientiousness, for which the cross-sectional correlation was $r_{true} = .30$, but longitudinal $r_{true} = .12$; however, the earlier Big Five scores had similar cross-sectional and longitudinal correlations with LS. At the item level, later personality traits' correlations with earlier and later LS were highly similar, with the two vectors of item–LS correlations (z-transformed) correlating .91. At the facet level, the earlier personality traits' correlations with LS were partly driven by four facets: N3 depression, N6 vulnerability, E6 positive emotions, and C1 competence (Supplemental Table S9). We also predicted the earlier LS from models trained to predict the later LS from the later-assessed 136 personality items and Big Five domains, omitting participants with earlier data from model training; the respective true predictive accuracies were .75 (*SE* = .051) and .61 (*SE* = .058). So, individual differences in both personality traits and LS as well as their correlations tended to endure over time, and it did not matter much whether r_{true}s were calculated, and LS predicted, cross-sectionally and longitudinally.

Domain-Specific Life Satisfactions

Next, we cross-validated LS and its correlations with personality traits against individual DSs' and their aggregate, representing an alternative way of conceptualizing and assessing general satisfaction. In summary, the r_{true} between aggregate DS and LS was .87, suggesting that LS's assessments closely tracked how satisfied people were with several specific life domains combined; for reference, self-reported LS and DS aggregate correlated $r = .67$. Likewise, LS was linked with all DS items, especially those referring to satisfaction with relationships and financial situation ($r_{true} > .65$; Supplemental Table S10). Moreover, the different DSs' correlations with informant-reported personality items were similar (vector correlation *Mdn* = .81), and the DS and LS aggregates' respective correlations with personality items were nearly identical, with a vector correlation of $r = .98$; see Supplemental Analyses 2 for details. Finally, individual DSs and their aggregate could be predicted from informant-reported personality traits with accuracies comparable to predicting LS from these same traits. In fact, the model trained to predict LS predicted the DS aggregate almost as well, and it also predicted all individual DSs; see Supplemental Analyses 2 for details.

So, DSs' overall extents and details of reflecting a broad range of personality traits were fairly similar among themselves and with LS, providing strong evidence for the robustness of the findings across different ways of assessing satisfaction.

Replications in Russian and English

The patterns of findings among Estonian speakers replicated well in smaller Russian- and English-speaking samples, allowing for some sampling variance in these comparatively smaller samples; see Supplemental Analyses 3 for further details. As one of the key findings, the vectors of LS aggregate's correlations with personality items were highly similar in the three samples. We meta-analytically combined the r_{true}s in Russian- and English-based data for the 69 personality items common to both samples, which correlated $r = .97$ with the respective Estonian-based r_{true}s, even though the 99% confidence intervals of these meta-analytic r_{true}s did not span the respective Estonian-based estimates for 12 items (Figure 1). As in Estonian-based data, openness and agreeableness were less correlated with LS than other domains among Russian

[6] The two directions of predicting LS from personality traits yielded similar results. For example, informant-rated LS correlated with its values predicted from the 136 self-rated items was .47, while self-rated LS's correlation with its values predicted from 136 informant-rated items was .45. So, there was no evidence that either informant- or self-rated LS would be more or less informative in relation to other traits.

[7] The earlier single-item LS correlated with the later scores of the most similar single item, "Am happy with my life," $r_{true} = .72$.

and English speakers, with $|r_{\text{true}}\text{s}|$ varying from 0 to .10 (Table 2). Emotional stability, extraversion, and conscientiousness had r_{true}s with the LS aggregate from .36 to .50 among Russian speakers and from .26 to .45 among English speakers, with their meta-analytic correlations being $r_{\text{true}} = .34$ (emotional stability), .47 (extraversion), and .38 (conscientiousness). Using elastic net models trained in the Estonian-based data to predict LS in Russian- and English-based data, the true predictive accuracies were .90 and .84 for items, respectively, and .74 and .64 for domains (Table 3). The findings pertaining to DSs also replicated well among people tested in Russian; data for most DSs were not available in English, and hence we did not replicate DS-related analyses in those data. For example, the individual DSs were correlated with LS similarly to the Estonian data (Supplemental Table S10), the LS–DS aggregates' r_{true} was .92 (Supplemental Table S10), and the predictive models trained in the Estonian-speakers' data were almost as predictive in the Russian-speakers' data (Supplemental Analyses 2).

So, LS's and DSs' true degrees of reflecting other personality traits, net of biases and random error, were strikingly robust across samples and languages.

Further Robustness Analyses

To address concerns raised during the articles' review process, we carried out three more robustness checks that are fully described in the Supplemental Analyses 4–6. First, we reduced LS to a single item, "Am happy with my life," instead of an aggregate of three items. This somewhat lowered the nuances' and domains' correlations with LS (e.g., among Estonian speakers, $r_{\text{true}} = .41$, .34, and .23, respectively, for emotional stability, extraversion, and conscientiousness, and prediction accuracy up to .82), but the overall patterns of findings remained highly similar in all three samples. Second, we addressed the possibility that LS could only be validly assessed using self-reports, and that the LS's cross-rater correlations only arose because informants had observed the targets' personality traits, which were correlated with the targets' otherwise private LS. Specifically, we removed LS's informant-reports from our calculations of true associations and predictive accuracies. Thus, r_{true}s were calculated as follows, with x representing the personality trait in question:

$$r_{\text{true}} = \frac{r_{\text{LS(self)}x\text{(informant)}}}{r_{x\text{(self)}x\text{(informant)}}}. \quad (4)$$

Again, this somewhat changed the results, but the overall patterns of findings remained highly similar in all three samples (e.g., among Estonian speakers, $r_{\text{true}} = .38$, .44, and .30, respectively, for emotional stability, extraversion, and conscientiousness, and prediction accuracy up to .83).

Finally, we used an alternative strategy for removing comparatively more overlapping personality assessment items, removing the item with the stronger (as opposed to weaker) LS correlations from each highly correlating personality item pair; see Supplemental Analyses 6 for details. This resulted in emotional stability having a somewhat stronger but extraversion and conscientiousness somewhat weaker r_{true} with LS and openness and agreeableness still being virtually unrelated to LS, for example, the respective correlations were .55, .39, .19, −.02, and .01 among the Estonian speakers. The alternative item retention strategy did not lower the domains and items true predictive accuracy for LS. Thus, although there is no inherent reason to prefer one strategy over the other, the strategy choice did not matter for our main conclusions.

Discussion

In one of the most comprehensive studies on this topic yet, we analyzed data from and across three samples where a range of personality traits and LS were rated by participants themselves and their informants. This allowed us to estimate LS's and personality traits' true associations free of single-method biases, occasion-specific effects, and random error. Besides avoiding direct construct overlaps at the item level, we cross-validated the findings with a different way of assessing satisfaction: an aggregate of satisfactions with eight specific life domains (DSs). Our findings suggest that in a world without common yet usually unaddressed measurement limitations, it would be possible to fairly accurately predict someone's satisfaction from a handful of personality traits. Specifically, correlations between actual LS and its values predicted from the Big Five personality domains or nuances could reach about .90, even when the predictions were based on associations found in an independent sample tested in a different language. Strikingly, even just three personality items allowed us to predict LS with approximately .80 accuracy. Moreover, LS could be predicted from personality traits with around .70 accuracy over approximately 10 years, similarly to LS's own stability.

We had no reason to a priori expect such findings. Because associations observed in typical single-method studies are likely inflated by shared method biases and at least sometimes by trivial construct overlaps, we could have found that LS's true predictability is much lower than is usually observed.[8] Yet, the predictability of LS turned out to be considerably higher. It expected that LS overlaps with personality traits to some degree in normal circumstances, where people can shape and evaluate their life according to their traits. But our estimates of this overlap's true extent are strikingly high, suggesting that how satisfied people are with their lives is usually quite close to what one could expect from their personality traits (Costa & McCrae, 1980). So, most life circumstances and other influences that are relevant for LS are those that also shape personality traits more broadly and endure over time.

How Much Higher Than the Usual Estimates?

Depending on the questionnaire, the Big Five domains have explained about 30% of LS's variance in self-report studies (Anglim et al., 2020; Busseri & Erb, 2023). This translates to a maximum out-of-sample predictive accuracy of .55, optimistically assuming no overfitting (Yarkoni & Westfall, 2017). Comprehensive facets sets such as those of the NEO Personality Inventory (Costa & McCrae, 1992) may explain up to about 40% of LS's variance (Anglim et al., 2020), translating to a maximum out-of-sample prediction of approximately .65. But these facets' assessments often directly ask about hopelessness, worthlessness, happiness, and optimism, and may therefore suffer from construct overlaps with LS, potentially leading to its overestimated predictability. Indeed, less expansive facet sets explain less LS variance. For example, in a large sample tested with the revised Big Five Inventory–2 (Soto & John, 2017),

[8] Our informal conversations with personality/LS researchers before writing this article reinforced precisely that expectation.

Stewart et al. (2022) found out-of-sample predictive accuracies of .48 and .50 for LS, respectively, for the Big Five domains and facets. In our single-method data, the Big Five domains provided about .65 out-of-sample predictive accuracy for LS. This suggests that our Big Five scales inherently captured more LS-related variance than many other Big Five scales, despite avoiding direct item overlaps. One plausible reason is that our Big Five scores were nearly orthogonal, thus capturing more personality variance in aggregate.

Generously putting the usual estimates' higher bound at .65, this is less than three quarters of the nearly .80 true predictive accuracy we observed for the Big Five domains and just over half of the .90 true predictive accuracy for items that capture personality nuances besides domains (after having z-transformed the correlations to make such comparisons meaningful due to correlations' nonlinear scale). Therefore, the extent to which LS reflects personality traits may be underestimated by a factor of nearly two in typical single-method Big Five studies, including various meta-analyses (e.g., Anglim et al., 2020). But again, our findings would have been equally meaningful even if the findings did resemble typical estimates of single-method studies because these could have been biased upward or downward, and this could not have been known a priori.

Even regardless of how individual researchers prefer to theorize on the personality trait–LS overlap, the mere fact that this overlap may be about twice as strong as typical findings show is highly important in and of itself and must constrain any theorizing on LS's origins. That researchers care about this overlap's degree is evidenced by the thousands of citations to previous meta-analyses such as DeNeve and Cooper (1998) and Steel et al. (2008) and the hundreds of citations already attracted by Anglim et al. (2020), despite the results of these meta-analyses being likely distorted due to unaddressed measurement issues.

Not Just Semantically Overlapping Evaluations

It is possible that people's general evaluations of their lives (e.g., "Am happy with my life") partly overlap with their personality trait evaluations (e.g., "Am energetic," "Often feel misunderstood") for reasons that are trivial or make LS's assessments inconsistent with its definition. For example, the items may appear semantically overlapping, or people may (consciously or unconsciously) think about their personality rather than their life per se when assessing their LS. However, assuming that people's evaluations of various specific life domains such as their job, career choice, relationships, financial situation, health, appearance, home, and country are less likely to overlap with these same personality traits for these same reasons, our results circumvent the possibility that LS's associations with personality traits are trivial.[9] This is because LS was highly correlated with a combination of eight diverse DSs, and the different DSs largely shared (informant-reported) personality correlates among themselves and with LS. Moreover, the model trained to predict LS allowed predicting the DS aggregate almost as accurately, besides allowing to predict each individual DS. Also, we ensured that no personality item correlated with LS items more strongly than LS items correlated among themselves, supporting the traits' discriminant validity.

Robustness Across Samples and Languages

It is also reasonable to think that LS's meaning and correlations with personality traits may be sensitive to context and/or assessment language; thus, not necessarily replicating across diverse samples. Also, other factors may influence LS to different degrees across samples, leaving more or less room for personality-related variance. If so, for example, even findings based on the whole Estonian population would have limited relevance for the French, Americans, Angolans, or Vietnamese. Indeed, there is already evidence that LS's correlations with positive and negative affect can systematically vary across countries (Kööts-Ausmees et al., 2013; Kuppens et al., 2008).

In our data, some true correlations, r_{true}s, did indeed vary across samples tested in different languages. For example, emotional stability's correlation with LS was higher among Estonian speakers ($r_{\text{true}} = .47$) than among those tested in Russian/English ($r_{\text{true}} = .34$), whereas the correlation with conscientiousness was lower ($r_{\text{true}} = .30$ vs. $r_{\text{true}} = .38$). Several individual items, reflecting partly unique personality nuances within and beyond the Big Five domains, also had somewhat different correlations among Estonian speakers than in other samples. However, although these cross-sample differences could speak to important questions about LS's context-sensitivity (besides possible slight translation differences), here we focus on the big picture according to our data: the patterns of how LS and DSs were related to one another and a range of personality traits remained highly replicable across three samples tested in different languages. This is best illustrated by our finding that models trained to predict LS in the Estonian-speaking sample tended to about as accurately predict LS among people tested in Russian (living in Estonia) and English (mostly living in Western Europe). This would not have been possible if LS's personality correlates were highly contextual. Such cross-sample predictive accuracy also has methodological implications, making it unlikely that the models were overfitted to data and that more complex models' predictive advantages reflected model complexity (Mõttus et al., 2020).

This does not mean that LS's true associations with DSs and personality traits could not vary across more diverse samples, such as those with non-European backgrounds or living in vastly different socioeconomic circumstances. This could be tested in future research, for which our methods can offer a blueprint and our results can offer a benchmark.

LS's Stable Variance Is Largely Shared With Personality Traits

Like personality traits, LS is far from perfectly stable over time. But personality traits' involvement in LS appears to endure over time because their longitudinal associations over several years were about as strong as cross-sectional ones. In fact, people's LS about 10 years earlier could be predicted from their later personality traits at least as accurately as it could be predicted from LS itself. So, like personality traits, LS may fluctuate spontaneously or respond to variable circumstances, but its stable variance is largely shared with that of personality traits. This finding also mitigates the concern that our findings may be specific to circumstances concurrent to our main data collection, such as the pandemic or the then-looming Russian invasion of Ukraine.

[9] For example, even if people (partly) base their rating of the item "Am happy with my life" on how well they think others understand them (or these items are semantically overlapping), this seems less likely for items asking about satisfaction with health, appearance, and residency.

Is Satisfaction Just a Reflection of Other Traits?

Although different ways of assessing satisfaction—LS and DSs—strongly overlap with personality traits in usual circumstances, being satisfied (with life) can remain conceptually distinct from personality more broadly. In some hypothetical circumstances, the associations of LS and DSs with some personality traits, such as feeling understood by others, might be weakened because the satisfactions are primarily shaped by strong external influences beyond an individual's control, while the personality traits may remain less influenced. As one possibility, thus, the strength of the overlap between LS and personality traits can be seen as a measure of the extent to which individuals can influence and assess their lives according to their traits. The more satisfaction appears as a stable, observable, and partly heritable trait that similarly manifests across different life domains and is entangled with other traits, the more it could reflect people's own choices, aspirations, behaviors, skills, and emotional and cognitive processes, rather than external circumstances imposed on them without their own involvement. While here this remains an untested hypothesis meant to illustrate the conceptual distinction between LS and personality traits, it could be tested by studying personality trait–LS associations in highly unusual, uncontrollable, and restrictive circumstances such as living in a war zone (for relevant studies, see Cheung et al., 2020 and Coupe & Obrizan, 2016).

It is also unnecessary to assume that particular personality traits are LS's directly interpretable causes even when they strongly correlate with LS. People differ in many traits, and each of these can contribute to, and be further shaped by, multiple traits and outcomes, including LS. This means that causal contributions can crisscross multiple traits and outcomes in any number of ways (Avinun, 2020), making them correlated over time but potentially leaving some or many of the individual pathways too complex to be meaningfully interpretable on their own (Brown & Rohrer, 2020; Mõttus et al., 2020). If so, the overall correlatedness among personality traits and variables like LS might often provide as much, if not more, insight as their individual associations.

Still Room for Other Influences

Although correlations as high as .90 are uncommon in psychology, even when corrected for measurement error, they must not be overinterpreted. Even such strong population trends leave considerable room for individuals to deviate from the statistical expectations, especially for those with the variables' medium levels (Mõttus, 2022). For example, if we trisected both predicted and observed LS, their .90 correlation would mean that the predicted and actual LS levels are different for every fourth individual. More specifically, every fifth individual predicted to have a high or low LS would actually have a different LS level, whereas among those predicted to have a medium LS, nearly two out of five would defy the prediction (Figure 2). Put differently, as the typical difference between two normally distributed measurements correlating at .90 is approximately a third of a standard deviation, most individuals' observed LS differs from its predicted-from-personality value by about the influence one would expect from a consequential life event (e.g., Denissen et al., 2019; Luhmann et al., 2012). This means there is still room for factors beyond those also captured personality traits to explain why some people's LS is higher or lower than expected from their personality traits. However, the factors that are also captured in personality traits—enduring life circumstances, idiosyncratic experiences, or genetics—still matter relatively more for most people, most of the time. So, our findings do not negate but constrain theories that aim to explain LS with factors completely external to the individual.

Which Traits Are Most Strongly Linked With LS?

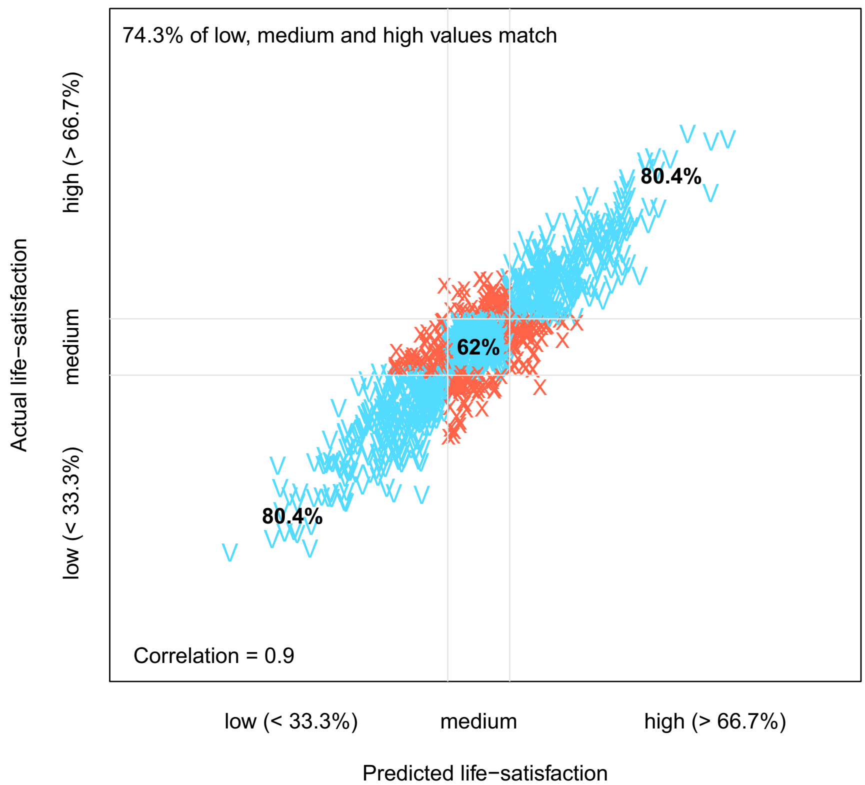
For the Big Five domains, LS had .30–.50 r_{true}s with emotional stability, extraversion, and conscientiousness. Although somewhat higher, these correlations are consistent with those usually observed in single-method studies (Anglim et al., 2020; Table 2), despite not being influenced by several common methodological limitations of these studies that could substantially deflate or inflate their findings. However, agreeableness and openness had small r_{true}s with LS, while agreeableness usually has stronger correlations with LS in single-method studies (see Anglim et al., 2020). Given our single-method correlations, this difference from previous research could partly result from our use of r_{true}s and partly from the near-orthogonality of our Big Five domains; undesirably, the domains are usually more intercorrelated in other Big Five measures, contributing to spurious correlations with other variables (Busseri & Erb, 2023; Stewart et al., 2022). Previous work with orthogonal Big Five scores has also resulted in somewhat weaker agreeableness–LS correlations, at least in self-reports (Busseri & Erb, 2023; McCrae & Costa, 1991). However, another part of the explanation may lie with cultural differences because the NEO-PI-3 agreeableness domain also had relatively small r_{true}s and single-method correlations with LS in the Estonian-speaking data (Supplemental Table S9); the r_{true} was also slightly higher among our English-speaking participants.

However, the Big Five items often differed in their r_{true}s with LS, as is common for many other outcomes (e.g., Revelle et al., 2021; Seeboth & Mõttus, 2018; Stewart et al., 2022). Many items also had stronger correlations with LS than any domain, including items not included in the domains. As for LS's strongest and relatively distinct correlates, its low levels were associated with feeling misunderstood, unexcited, indecisive, envious, bored, and used by others, whereas high LS tended to go with confidence in one's abilities and believing that efforts are rewarded.

In fact, even three items, respectively, reflecting feeling misunderstood, lack of excitement, and being indecisive, provided (out-of-sample) prediction of LS with true accuracy of .80, suggesting that most people with low LS could be recognized from just a few personality nuances. This is comparable to the predictive accuracy provided by the Big Five domains that encompass a broad range of traits each, some of which are more and some less correlated with LS, making the domain-level results more ambiguous (Mõttus, 2016) and conducive to "just so" stories. For example, if we were only told that many people with low LS are low on emotional stability, extraversion, and conscientiousness, we could explain low LS by referring to any number of traits subsumed under these broad domains, only a few of which might actually correlate with LS. So, without further specifics, we could easily indulge in baseless speculations. Instead, now knowing that most people with low LS tend to feel misunderstood, lack excitement, and struggle with making decisions, our degrees of freedom in explaining low LS become smaller. Also, low emotional stability, low extraversion, and

Figure 2

Predicted-From-Personality and Actual Satisfaction Levels, Overlapping for Blue Individuals (74%) but Being Incongruent for Red Individuals (26%)



Note. See the online article for the color version of this figure.

low conscientiousness tend to go with many undesirable outcomes, offering limited discriminant validity in explaining these outcomes, although these outcomes often also go with lower openness and agreeableness (Bleidorn et al., 2020) which do not have nontrivial r_{true}s with LS. It remains to be seen if LS's more nuanced r_{true}s with personality traits are specific to this outcome, offering greater discriminant validity and suggesting that there are factors that shape LS specifically rather than a desirable life more generally.

In conclusion, although the Big Few domains will continue to provide a parsimonious representation of LS's associations with personality traits, supplementing domain-level analyses with nuances offers a richer and more accurate picture of how LS intersects with psychological traits more broadly, besides providing greater predictive accuracy. The ability to estimate error-free associations with multirater or multitimepoint data (Wood et al., 2023) is particularly useful for this research because it makes nuance-level associations directly comparable to those of aggregate personality traits.

The Self Is Not Privileged to Evaluate LS

One may think that people's LS levels are private. However, if we accept that people's personality traits are to some extent observable to others (Connelly & Ones, 2010), then we have to accept the same for LS because its cross-rater agreement is similar to that of personality traits (Dobewall et al., 2013; Schneider & Schimmack, 2009). Moreover, we found that self-reported personality traits' correlations with informant-reported LS were similar to informant-reported personality traits' correlations with self-reported LS (see Footnotes 4 to 6), suggesting that self- and informant-reports of LS contained comparable degrees of information about personality traits. Besides, we calculated domains' and nuances' r_{true}s with LS and their true predictive accuracies for LS based on just self-reported LS (Supplemental Analyses), and none of the findings were different enough to change our conclusions, further alleviating the concerns that our findings could have been an artifact of combining self-reported LS with informant-reported LS. Further, our DS-related analyses did not include informant-reports, yet the patterns of findings were similar to those of LS-related analyses that did include informant-reports.

Of course, self-informant agreement is high for neither LS nor personality traits. For example, if we trisected self- and informant-report scores that correlate about .50, the targets' scores would be similar in only about half of the self-informant pairs (Mõttus, 2022). However, such moderate self-informant agreement is the very reason why our findings are particularly novel and meaningful, showing that much of trait scores' variance is specific to a single method which can strongly bias the traits' correlations among themselves and with other variables in typical single-method data. Besides, our method only required that there was some agreement and, ideally, that self- and informant-reports were available for both variables being correlated; the imperfect agreement would then cancel out because it would similarly influence both the numerator and denominator in the r_{true} calculations (see Supplemental Material for the algebraic proof). For our analyses involving DSs, informant-reports were unavailable, so their r_{true}s could have been somewhat distorted. However, given that there was a substantial level of cross-rater agreement for all items for which both self- and informant-reports were available—personality items, LS items, and three DSs items—and DS-related and LS-related findings were similar, it is unlikely that even the DS-related r_{true}s were distorted enough to bias our conclusions.

In short, we found no compelling evidence that our use of informant-reported LS, in addition to self-reports, caused the observed pattern of findings.

LS's Construct Validity

Desirably, the three LS items assessed slightly different aspects of the construct because their r_{true}s were around .75 in the Estonian-speaking sample, unlike the near-unity r_{true}s among semantically nearly identical personality items before we removed redundant items. In the smaller samples tested in Russian and English, the three items had somewhat more variable but still high r_{true}s among themselves, with the variability likely due to their higher sampling variance (all correlations were within ±2 *SE* from the Estonian estimates). Substantially higher r_{true}s among the LS items would have been undesirable, narrowing the construct's scope (Clark & Watson, 1995). Supporting LS's construct validity, its items correlated among themselves more strongly than they correlated with personality items, and they had highly similar correlation profiles with personality items, showing similar broader psychological backgrounds. In the Supplemental Analyses, we also showed that LS's r_{true}s with domains and nuances, as well as personality traits' true predictive accuracy for LS, would have been quite similar—although generally somewhat lower due to LS being more narrowly defined—if we had assessed LS with only one single item, "Am happy with my life." Moreover, in Supplemental Analyses, we also showed that our LS assessments correlated very highly—r = .80 (in Russian), .90 (in English), and .95 (in Estonian)—with the widely used SWLS scale (Diener et al., 1985), and it had a high 10-year longitudinal correlation with another LS's assessment. Finally, LS correlated very highly with the aggregate of a range of DSs, which in turn correlated extremely highly with the SWLS, further aligning the LS's assessment with its definition. Thus, although our LS assessment could be seen as unconventional by some, it provided a very close proxy measure for LS.

Our three LS items covered a general life satisfaction assessment ("Am happy with my life"), purpose in life ("Feel that my life lacks direction"), and perspective on the future ("Have a dark outlook on the future"). Arguably, thus, our LS assessment had a broader scope than the SWLS (Diener et al., 1985) despite their very high empirical overlap. For example, our LS assessment also covered an aspect of the eudaimonic well-being (purpose) besides the hedonic well-being aspects usually associated with LS (Ryff et al., 2021).[10] Given this, it is not surprising that the LS's strongest correlates included items beyond those directly referring to emotional well-being. In particular, our LS assessment and many of its correlates fit with the components of Ryff's (1989) model of psychological well-being, which includes positive relations (e.g., items about feeling understood, trust, liking others, and enjoying cooperation), autonomy, environmental mastery and personal growth (e.g., items about self-competence, learning quickly, solving complex problems, leadership and influencing others, believing in hard work, taking risks, learning new things, visiting new places, and self-improvement), purpose in life (e.g., items about lack of excitement, indecisiveness, avoiding responsibilities, and boredom) and self-acceptance (e.g., an item about wisdom). Thus, the hedonic and eudaimonic well-being aspects may overlap more than sometimes thought, both empirically and in their broader psychological correlates. This is also consistent with previous work on the hedonic and eudaimonic well-being overlap, based on self-reports alone (e.g., Disabato et al., 2016).

In conclusion, we believe that our findings provide strong evidence for the validity of a broad LS construct in general and our LS assessments in particular.

Limitations

Our study did not have the usual limitations of personality research, such as relying on a monocultural sample, self-report-only measures, brief questionnaires, broad trait domains, or a single operationalization of the target construct, nor did it suffer from limited statistical power. However, although our personality item pool was intentionally expansive, it almost certainly did not cover all possible personality nuances; hence, likely missing some LS-relevant personality traits. If so, we could underestimate personality traits' predictive accuracy for LS. However, given that our estimated true predictive accuracy was as high as .90, the completely missed personality content could not have been extensive. Likewise, our list of eight DSs likely missed some life domains that may be particularly relevant to some people's well-being. Also, the DSs were only assessed with self-reports, introducing possible biases to their r_{true}s with LS and personality traits. Further, our samples were convenience samples with a high percentage of females and high levels of education, possibly leading to underestimated correlations due to reduced variance. Future studies should aim to generalize our findings to more diverse populations.

Conclusion

When addressing common methodological limitations, most people's LS levels are accurately predictable from their personality traits, even when avoiding direct construct overlaps. This does not mean LS is inherently and irrevocably reducible to personality traits. Instead, the degree to which it reflects personality traits may be considered a measure of people's freedom to shape and assess their lives according to their traits. At least hypothetically, there could be circumstances where LS is less aligned with personality traits. But in usual conditions, there does not seem to be much reason to think that LS is mostly shaped by circumstances unrelated to personality more broadly. For most people, and most of the time, their satisfaction level is just about what we would expect from their personality traits. Personality traits can be shaped by any number of factors, but

[10] We dropped the item referring to life having been kind to the person because it was less consistent with other LS items in all three languages. However, the SWLS has a parallel item: "So far I have gotten the important things I want in life."

usually, these same factors also shape LS, through personality or otherwise.

References

Anglim, J., Horwood, S., Smillie, L. D., Marrero, R. J., & Wood, J. K. (2020). Predicting psychological and subjective well-being from personality: A meta-analysis. *Psychological Bulletin*, *146*(4), 279–323. https://doi.org/10.1037/bul0000226

Ashton, M. C., & Lee, K. (2020). Objections to the HEXACO model of personality structure—And why those objections fail. *European Journal of Personality*, *34*(4), 492–510. https://doi.org/10.1002/per.2242

Avinun, R. (2020). The E is in the G: Gene–environment-trait correlations and findings from genome-wide association studies. *Perspectives on Psychological Science*, *15*(1), 81–89. https://doi.org/10.1177/1745691619867107

Bainbridge, T. F., Ludeke, S. G., & Smillie, L. D. (2022). Evaluating the Big Five as an organizing framework for commonly used psychological trait scales. *Journal of Personality and Social Psychology*, *122*(4), 749–777. https://doi.org/10.1037/pspp0000395

Beck, E. D., & Jackson, J. J. (2022). A mega-analysis of personality prediction: Robustness and boundary conditions. *Journal of Personality and Social Psychology*, *122*(3), 523–553. https://doi.org/10.1037/pspp0000386

Bleidorn, W., Hopwood, C. J., Ackerman, R. A., Witt, E. A., Kandler, C., Riemann, R., Samuel, D. B., & Donnellan, M. B. (2020). The healthy personality from a basic trait perspective. *Journal of Personality and Social Psychology*, *118*(6), 1207–1225. https://doi.org/10.1037/pspp0000231

Bouchard, T. J., Jr. (2016). Experience producing drive theory: Personality "writ large". *Personality and Individual Differences*, *90*, 302–314. https://doi.org/10.1016/j.paid.2015.11.007

Boyce, C. J., Brown, G. D. A., & Moore, S. C. (2010). Money and happiness: Rank of income, not income, affects life satisfaction. *Psychological Science*, *21*(4), 471–475. https://doi.org/10.1177/0956797610362671

Brown, N. J. L., & Rohrer, J. M. (2020). Easy as (happiness) pie? A critical evaluation of a popular model of the determinants of well-being. *Journal of Happiness Studies*, *21*(4), 1285–1301. https://doi.org/10.1007/s10902-019-00128-4

Busseri, M. A., & Erb, E. M. (2023). The happy personality revisited: Re-examining associations between Big Five personality traits and subjective well-being using meta-analytic structural equation modeling. *Journal of Personality*. Advance online publication. https://doi.org/10.1111/jopy.12862

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait–multimethod matrix. *Psychological Bulletin*, *56*(2), 81–105. https://doi.org/10.1037/h0046016

Caspi, A., Roberts, B. W., & Shiner, R. L. (2005). Personality development: Stability and change. *Annual Review of Psychology*, *56*(1), 453–484. https://doi.org/10.1146/annurev.psych.55.090902.141913

Cheung, F., Kube, A., Tay, L., Diener, E., Jackson, J. J., Lucas, R. E., Ni, M. Y., & Leung, G. M. (2020). The impact of the Syrian conflict on population well-being. *Nature Communications*, *11*(1), Article 3899. https://doi.org/10.1038/s41467-020-17369-0

Cheung, F., & Lucas, R. E. (2016). Income inequality is associated with stronger social comparison effects: The effect of relative income on life satisfaction. *Journal of Personality and Social Psychology*, *110*(2), 332–341. https://doi.org/10.1037/pspp0000059

Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, *7*(3), 309–319. https://doi.org/10.1037/1040-3590.7.3.309

Condon, D. M., & Revelle, W. (2016). Selected ICAR data from the SAPA-project: Development and initial validation of a public-domain measure. *Journal of Open Psychology Data*, *4*(1), Article e1. https://doi.org/10.5334/jopd.25

Condon, D. M., Wood, D., Mõttus, R., Booth, T., Costantini, G., Greiff, S., Johnson, W., Lukaszewski, A., Murray, A., Revelle, W., Wright, A. G. C., Ziegler, M., & Zimmermann, J. (2020). Bottom up construction of a personality taxonomy. *European Journal of Psychological Assessment*, *36*(6), 923–934. https://doi.org/10.1027/1015-5759/a000626

Connelly, B. S., & Ones, D. S. (2010). An other perspective on personality: Meta-analytic integration of observers' accuracy and predictive validity. *Psychological Bulletin*, *136*(6), 1092–1122. https://doi.org/10.1037/a0021212

Costa, P. T., Jr., & McCrae, R. R. (1980). Influence of extraversion and neuroticism on subjective well-being: Happy and unhappy people. *Journal of Personality and Social Psychology*, *38*(4), 668–678. https://doi.org/10.1037/0022-3514.38.4.668

Costa, P. T., Jr., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual*. Psychological Assessment Resources.

Coupe, T., & Obrizan, M. (2016). The impact of war on happiness: The case of Ukraine. *Journal of Economic Behavior & Organization*, *132*(Pt. A), 228–242. https://doi.org/10.1016/j.jebo.2016.09.017

DeNeve, K. M., & Cooper, H. (1998). The happy personality: A meta-analysis of 137 personality traits and subjective well-being. *Psychological Bulletin*, *124*(2), 197–229. https://doi.org/10.1037/0033-2909.124.2.197

Denissen, J. J. A., Luhmann, M., Chung, J. M., & Bleidorn, W. (2019). Transactions between life events and personality traits across the adult lifespan. *Journal of Personality and Social Psychology*, *116*(4), 612–633. https://doi.org/10.1037/pspp0000196

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., Oishi, S., & Tay, L. (2018). Advances in subjective well-being research. *Nature Human Behaviour*, *2*(4), 253–260. https://doi.org/10.1038/s41562-018-0307-6

Diener, E., & Seligman, M. E. P. (2002). Very happy people. *Psychological Science*, *13*(1), 81–84. https://doi.org/10.1111/1467-9280.00415

Disabato, D. J., Goodman, F. R., Kashdan, T. B., Short, J. L., & Jarden, A. (2016). Different types of well-being? A cross-cultural examination of hedonic and eudaimonic well-being. *Psychological Assessment*, *28*(5), 471–482. https://doi.org/10.1037/pas0000209

Dobewall, H., Realo, A., Allik, J., Esko, T., & Metspalu, A. (2013). Self-other agreement in happiness and life-satisfaction: The role of personality traits. *Social Indicators Research*, *114*(2), 479–492. https://doi.org/10.1007/s11205-012-0157-y

Funder, D. C., & Ozer, D. J. (2019). Evaluating effect size in psychological research: Sense and nonsense. *Advances in Methods and Practices in Psychological Science*, *2*(2), 156–168. https://doi.org/10.1177/2515245919847202

Goldberg, L. R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. In I. Mervielde, I. Deary, F. De Fruyt, & F. Ostendorf (Eds.), *Personality psychology in Europe* (Vol. 7, pp. 7–28). Tilburg University Press.

Heller, D., Watson, D., & Hies, R. (2004). The role of person versus situation in life satisfaction: A critical examination. *Psychological Bulletin*, *130*(4), 574–600. https://doi.org/10.1037/0033-2909.130.4.574

Henry, S., & Mõttus, R. (2022). *The 100 Nuances of personality: Development of a comprehensive, non-redundant personality item pool*. https://doi.org/10.17605/OSF.IO/TCFGZ

Jagodzinski, W. (2010). Economic, social, and cultural determinants of life satisfaction: Are there differences between Asia and Europe? *Social Indicators Research*, *97*(1), 85–104. https://doi.org/10.1007/s11205-009-9555-1

Johnson, W. (2010). Extending and testing Tom Bouchard's experience producing drive theory. *Personality and Individual Differences*, *49*(4), 296–301. https://doi.org/10.1016/j.paid.2009.11.022

Kööts-Ausmees, L., Realo, A., & Allik, J. (2013). The relationship between life satisfaction and emotional experience in 21 European countries. *Journal of Cross-Cultural Psychology*, *44*(2), 223–244. https://doi.org/10.1177/0022022112451054

Kuppens, P., Realo, A., & Diener, E. (2008). The role of positive and negative emotions in life satisfaction judgment across nations. *Journal of Personality and Social Psychology*, *95*(1), 66–75. https://doi.org/10.1037/0022-3514.95.1.66

Lucas, R. E., Freedman, V. A., & Cornman, J. C. (2018). The short-term stability of life satisfaction judgments. *Emotion*, *18*(7), 1024–1031. https://doi.org/10.1037/emo0000357

Luhmann, M., Hofmann, W., Eid, M., & Lucas, R. E. (2012). Subjective well-being and adaptation to life events: A meta-analysis. *Journal of Personality and Social Psychology*, *102*(3), 592–615. https://doi.org/10.1037/a0025948

McCrae, R. R. (2015). A more nuanced view of reliability: Specificity in the trait hierarchy. *Personality and Social Psychology Review*, *19*(2), 97–112. https://doi.org/10.1177/1088868314541857

McCrae, R. R., & Costa, P. T., Jr. (1991). Adding liebe und arbeit: The full five-factor model and well-being. *Personality and Social Psychology Bulletin*, *17*(2), 227–232. https://doi.org/10.1177/014616729101700217

McCrae, R. R., Costa, P. T., Jr., & Martin, T. A. (2005). The NEO-PI-3: A more readable revised NEO Personality Inventory. *Journal of Personality Assessment*, *84*(3), 261–270. https://doi.org/10.1207/s15327752jpa8403_05

McCrae, R. R., & Mõttus, R. (2019). What personality scales measure: A new psychometrics and its implications for theory and assessment. *Current Directions in Psychological Science*, *28*(4), 415–420. https://doi.org/10.1177/0963721419849559

McMahon, D. M. (2006). *Happiness: A history*. Gove Press.

Mõttus, R. (2016). Towards more rigorous personality trait-outcome research. *European Journal of Personality*, *30*(4), 292–303. https://doi.org/10.1002/per.2041

Mõttus, R. (2022). What correlations mean for individual people: A tutorial for researchers, students and the public. *Personality Science*, *3*, Article e7467. https://doi.org/10.5964/ps.7467

Mõttus, R. (2023, December 14). *Personality traits and life-satisfacton*. Open Science Framework. https://doi.org/10.17605/OSF.IO/YW7X3

Mõttus, R., Sinick, J., Terracciano, A., Hřebíčková, M., Kandler, C., Ando, J., Mortensen, E. L., Colodro-Conde, L., & Jang, K. L. (2019). Personality characteristics below facets: A replication and meta-analysis of cross-rater agreement, rank-order stability, heritability, and utility of personality nuances. *Journal of Personality and Social Psychology*, *117*(4), e35–e50. https://doi.org/10.1037/pspp0000202

Mõttus, R., Wood, D., Condon, D. M., Back, M. D., Baumert, A., Costantini, G., Epskamp, S., Greiff, S., Johnson, W., Lukaszewski, A., Murray, A., Revelle, W., Wright, A. G. C., Yarkoni, T., Ziegler, M., & Zimmermann, J. (2020). Descriptive, predictive and explanatory personality research: Different goals, different approaches, but a shared need to move beyond the big few traits. *European Journal of Personality*, *34*(6), 1175–1201. https://doi.org/10.1002/per.2311

Oishi, S., Diener, E. F., Lucas, R. E., & Suh, E. M. (1999). Cross-cultural variations in predictors of life satisfaction: Perspectives from needs and values. *Personality and Social Psychology Bulletin*, *25*(8), 980–990. https://doi.org/10.1177/01461672992511006

Paulhus, D. L., & Vazire, S. (2007). The self-report method. In R. W. Robins, R. C. Fraley, & R. F. Krueger (Eds.), *Handbook of research methods in personality psychology* (pp. 224–239). The Guilford Press.

Payne, J. W., & Schimmack, U. (2020). Construct validity of global life-satisfaction judgments: A look into the black box of self-informant agreement. *Journal of Research in Personality*, *89*, Article 104041. https://doi.org/10.1016/j.jrp.2020.104041

R Core Team. (2023). *R: A language and environment for statistical computing* [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org

Rentzsch, K., & Gross, J. J. (2015). Who turns green with envy? Conceptual and empirical perspectives on dispositional envy. *European Journal of Personality*, *29*(5), 530–547. https://doi.org/10.1002/per.2012

Revelle, W., Dworak, E. M., & Condon, D. M. (2021). Exploring the persome: The power of the item in understanding personality structure. *Personality and Individual Differences*, *169*, Article 109905. https://doi.org/10.1016/j.paid.2020.109905

Ryff, C. D. (1989). Happiness is everything, or is it? Explorations on the meaning of psychological well-being. *Journal of Personality and Social Psychology*, *57*(6), 1069–1081. https://doi.org/10.1037/0022-3514.57.6.1069

Ryff, C. D., Boylan, J. M., & Kirsch, J. A. (2021). Eudaimonic and hedonic well-being: An integrative perspective with linkages to sociodemographic factors and health. In M. T. Lee, L. D. Kubzansky, & T. J. Van der Weele (Eds.), *Measuring well-being: Interdisciplinary perspectives from the social sciences and the humanities* (pp. 92–135). Oxford University Press. https://doi.org/10.1093/oso/9780197512531.003.0005

Schimmack, U. (2010). What multi-method data tell us about construct validity. *European Journal of Personality*, *24*(3), 241–257. https://doi.org/10.1002/per.771

Schimmack, U., Oishi, S., Furr, R. M., & Funder, D. C. (2004). Personality and life satisfaction: A facet-level analysis. *Personality and Social Psychology Bulletin*, *30*(8), 1062–1075. https://doi.org/10.1177/0146167204264292

Schneider, L., & Schimmack, U. (2009). Self-informant agreement in well-being ratings: A meta-analysis. *Social Indicators Research*, *94*(3), 363–376. https://doi.org/10.1007/s11205-009-9440-y

Seeboth, A., & Mõttus, R. (2018). Successful explanations start with accurate descriptions: Questionnaire items as personality markers for more accurate predictions. *European Journal of Personality*, *32*(3), 186–201. https://doi.org/10.1002/per.2147

Soto, C. J. (2019). How replicable are links between personality traits and consequential life outcomes? The life outcomes of personality replication project. *Psychological Science*, *30*(5), 711–727. https://doi.org/10.1177/0956797619831612

Soto, C. J., & John, O. P. (2017). The next Big Five Inventory (BFI-2): Developing and assessing a hierarchical model with 15 facets to enhance bandwidth, fidelity, and predictive power. *Journal of Personality and Social Psychology*, *113*(1), 117–143. https://doi.org/10.1037/pspp0000096

Steel, P., Schmidt, J., & Shultz, J. (2008). Refining the relationship between personality and subjective well-being. *Psychological Bulletin*, *134*(1), 138–161. https://doi.org/10.1037/0033-2909.134.1.138

Stewart, R. D., Mõttus, R., Seeboth, A., Soto, C. J., & Johnson, W. (2022). The finer details? The predictability of life outcomes from Big Five domains, facets, and nuances. *Journal of Personality*, *90*(2), 167–182. https://doi.org/10.1111/jopy.12660

Suh, E., Diener, E., Oishi, S., & Triandis, H. C. (1998). The shifting basis of life satisfaction judgments across cultures: Emotions versus norms. *Journal of Personality and Social Psychology*, *74*(2), 482–493. https://doi.org/10.1037/0022-3514.74.2.482

Thielmann, I., Moshagen, M., Hilbig, B. E., & Zettler, I. (2022). On the comparability of basic personality models: Meta-analytic correspondence, scope, and orthogonality of the Big Five and HEXACO dimensions. *European Journal of Personality*, *36*(6), 870–900. https://doi.org/10.1177/08902070211026793

Turkheimer, E., Pettersson, E., & Horn, E. E. (2014). A phenotypic null hypothesis for the genetics of personality. *Annual Review of Psychology*, *65*(1), 515–540. https://doi.org/10.1146/annurev-psych-113011-143752

Vaht, M., Arumäe, K., Realo, A., Ausmees, L., Allik, J., Henry, S., Metspalu, A., Esko, T., Mõttus, R., & Vainik, U. (2024, February 24). *Cohort profile: Personality measurements at the Estonian Biobank of the Estonian Genome Center, University of Tartu*. https://doi.org/10.31234/osf.io/2aey6 van der Linden, D., te Nijenhuis, J., & Bakker, A. B. (2010). The general factor of personality: A meta-analysis of Big Five intercorrelations and a criterion-related validity study. *Journal of Research in Personality*, *44*(3), 315–327. https://doi.org/10.1016/j.jrp.2010.03.003

Vazire, S. (2006). Informant reports: A cheap, fast, and easy method for personality assessment. *Journal of Research in Personality*, *40*(5), 472–481. https://doi.org/10.1016/j.jrp.2005.03.003

Vihalemm, T., Juzefovičs, J., & Leppik, M. (2019). Identity and media-use strategies of the Estonian and Latvian Russian-speaking populations amid political crisis. *Europe-Asia Studies*, *71*(1), 48–70. https://doi.org/10.1080/09668136.2018.1533916

Weiss, A., Bates, T. C., & Luciano, M. (2008). Happiness is a personal(ity) thing: The genetics of personality and well-being in a representative sample. *Psychological Science*, *19*(3), 205–210. https://doi.org/10.1111/j.1467-9280.2008.02068.x

Wood, D. & Harms, P. D. (2016). On the TRAPs that make it dangerous to study personality with personality questionnaires. *European Journal of Personality*, *30*, 327–329. https://doi.org/10.1002/per.2060

Wood, D., Lowman, G. H., Armstrong, B. F., III, & Harms, P. D. (2023). Using retest-adjusted correlations as indicators of the semantic similarity of items. *Journal of Personality and Social Psychology*, *125*(2), 437–454. https://doi.org/10.1037/pspp0000441

Yarkoni, T., & Westfall, J. (2017). Choosing prediction over explanation in psychology: Lessons from machine learning. *Perspectives on Psychological Science*, *12*(6), 1100–1122. https://doi.org/10.1177/1745691617693393

Received January 8, 2023
Revision received December 14, 2023
Accepted February 1, 2024 ■