



# THE UNIVERSITY *of* EDINBURGH

## Edinburgh Research Explorer

### Towards more rigorous personality trait-outcome research

**Citation for published version:**

Mõttus, R 2016, 'Towards more rigorous personality trait-outcome research' *European Journal of Personality*, vol 30, no. 4, pp. 292-303. DOI: 10.1002/per.2041

**Digital Object Identifier (DOI):**

[10.1002/per.2041](https://doi.org/10.1002/per.2041)

**Link:**

[Link to publication record in Edinburgh Research Explorer](#)

**Document Version:**

Peer reviewed version

**Published In:**

*European Journal of Personality*

**General rights**

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



Running head: Personality traits and outcomes

**Towards more rigorous personality trait-outcome research**

René Mõttus \*

Department of Psychology, University of Edinburgh  
Institute of Psychology, University of Tartu

\* 8 George Square, EH8 9JZ Edinburgh, UK  
rene.mottus@ed.ac.uk

## **Abstract**

Much of personality research attempts to identify causal links between personality traits and various types of outcomes. I argue that causal interpretations require traits to be seen as existentially and holistically real and the associations to be independent of specific ways of operationalizing the traits. Among other things, this means that, to the extent that causality is to be ascribed to such holistic traits, items and facets of those traits should be similarly associated with specific outcomes, except for variability in the degrees to which they reflect the traits (i.e., factor loadings). I argue that, before drawing causal inferences about personality trait-outcome associations, presence of this condition should be routinely tested by, for example, systematically comparing the outcome associations of individual items or facets, or sampling different indicators for measuring the same purported traits. Existing evidence suggests that observed associations between personality traits and outcomes at least sometimes depend on which particular items or facets have been included in trait operationalizations, calling trait-level causal interpretations into question. However, this has rarely been considered in the literature. I argue that when outcome associations are specific to facets, they should not be generalized to traits. Furthermore, when the associations are specific to particular items, they should not even be generalized to facets.

Keywords: outcomes; facets; five-factor model; causality; ontology; realism

Psychologists routinely define and operationalize personality as a limited set of broad traits, each of which summarizes a wide range of thoughts, feelings and behaviors. As personality psychologists, we are interested in the internal structures and inter-relations of such traits, their genetic and environmental antecedents, developmental trends, neuroanatomical and physiological correlates, as well as their distributions across demographic groups, geographic locations and cultures. Here, I focus on another major line of research: the associations of these traits with various variables that presumably reflect something outside the domain of personality, to which I collectively refer as outcomes. These phenomena are often hypothesized to be partly caused by personality traits. Whether and when such causal interpretations are justified is the focus of this article.

### *Personality traits*

Among the broad personality trait taxonomies, I focus on the Five-Factor Model (FFM; McCrae & John, 1992) or Big Five (Goldberg, 1990) traits: Emotional Stability or Neuroticism, Extraversion, Openness to Experience or Intellect, Agreeableness, and Conscientiousness. This is not because I believe that these particular traits as commonly defined and operationalized necessarily define personality ‘as nature does’. These broad traits just happen to comprise the currently most popular and perhaps best-studied trait taxonomy, and studies linking personality characteristics to outcomes seem to have followed this trend. However, most of what I discuss equally applies to other broad trait taxonomies such as the PEN (Eysenck & Eysenck, 1985), MPQ (Tellegen & Waller, 2008) or HEXACO (Ashton & Lee, 2007) models.

### *Outcomes*

Broad personality traits are often linked to broad outcomes of particular social or personal importance, such as educational, occupational and marital success, health, happiness, and longevity (Ozer & Benet-Martínez, 2006; Roberts, Kuncel, Shiner, Caspi, & Goldberg, 2007). What these outcomes share is that, despite being quantified along one or a few dimensions, they are all complex

phenomena themselves. Besides factors that are generally considered unrelated to personality (e.g., parental education or country of birth), they likely reflect the cumulative and transactive effects of multiple behavioral choices that people make over long periods of time. This provides reasons to expect widespread links with personality trait scores, which also encompass wide ranges of behaviors, thoughts and feelings. Other types of outcomes refer to a variety of more specific behaviors such as volunteering (Carlo, Okun, Knight, & de Guzman, 2005), mate poaching (Schmitt & Shackelford, 2008), residential mobility (Jokela, 2009), eating habits (Möttus, McNeill, et al., 2013), smoking and drinking (Malouff, Thorsteinsson, Rooke, & Schutte, 2007; Malouff, Thorsteinsson, & Schutte, 2006), physical activity (Rhodes & Smith, 2006) or voting choices (Caprara, Schwartz, Capanna, Vecchione, & Barbaranelli, 2006), to name just a few.

#### *Why do we care?*

Broadly speaking, there are at least two reasons that researchers are interested in linking personality trait scores to outcomes. The first reason is to "empower" traits by demonstrating that they have important consequences (Roberts et al., 2007). The fact that personality trait scores are correlated with various types of outcomes is sometimes considered the main reason for interest in them (Goldberg, 1999; Ozer & Benet-Martínez, 2006). For example, an association between high scores on a personality trait measure and longevity may underscore the importance of studying the trait. The other main reason is to explain the variability in these outcomes. Among other things, it may be hoped that such associations, if sufficiently understood, can be put to practical use. For example, researchers may link personality traits to specific health conditions such as obesity or diabetes in the hope of understanding their risk factors, etiology or recurrence and thereby being able to devise better interventions or treatments for them (Deary, Weiss, & Batty, 2010).

#### *How are trait-outcome associations interpreted?*

Both of these motivations seem to require trait-outcome associations to be interpreted in causal terms. After all, particular personality traits such as Extraversion or Conscientiousness can

have consequences or provide explanations for phenomena only if they make causal contributions to them, at least probabilistically—in some people and/or in some circumstances. Although researchers are often careful to acknowledge *explicitly* that their study designs do not warrant causal inference, *implicitly* their interpretations often rely on it. For example, it has been suggested that increasing people's Conscientiousness might in principle help to increase their longevity, similarly to the effects of smoking reduction or increase in physical activity (Jokela et al., 2013). This suggestion, in fact coupled with a hypothetical point-estimate of mortality rate reduction, clearly entails the assumption that there are processes that, at least in some people and in some circumstances, link low Conscientiousness with exposure to more proximal contributors to death.

Here are some further examples of how researchers have proposed tentative causal interpretations, even though these are often qualified by alternative explanations or phrased cautiously. Möttus and colleagues (2013) discussed Openness as being “causal to healthy diet” (p. 358), although they also acknowledged the possibilities of causality running in the other direction or the correlation being spurious due to unmeasured variables. Malouff and colleagues (2007) suggested “causal links” (p. 277, 291) between personality traits and alcohol use, although admitting in another paragraph that the reviewed studies “do not provide the needed evidence to reach causal conclusions” (p. 291). In a similar vein, Roberts and colleagues (2007) showed in a comprehensive meta-analysis that high Emotional Stability, Agreeableness and Conscientiousness were associated with socio-economic success, lower probability of divorce, and longevity. Although they were “not assuming that personality traits are direct causes of the outcomes under study” (p. 315), they nevertheless used phrases referring to rather explicit causal interpretations throughout the paper (e.g., traits “lead to”, “affect”, “contribute”, or “exert their effects”). For instance, they suggested that “personality traits lead to divorce or conversely marital stability” because they “help shape the quality of long-term relationships” (p. 327). Jokela and colleagues (2014) concluded that “low conscientiousness is a potentially important upstream risk factor for developing diabetes” (p.

1620). Weston and Jackson (2015) argued that personality traits "may lead to" optimal levels of health behaviors such as smoking, drinking and exercise (p. 67), and Hagger-Johnson and colleagues (2011) claimed to have demonstrated that "personality [Conscientiousness] and individual differences are important *determinants* [italics mine] of both sexual health behaviors and alcohol use prior to and during sexual events" (p. 842).

To be clear, my intent is not to lament that researchers propose causal interpretations. Instead, I simply want to point out that researchers often seem to prefer or gravitate towards interpreting the personality trait-outcome associations they observed as causal, even when they—quite rightly—explicitly acknowledge this goes beyond a strict reading of the data. It is natural to strive towards causal explanations, given the reasons of doing this kind of research.

### **When are the causal interpretations warranted?**

Correlations are always difficult to interpret. Does  $x$  cause  $y$  or vice versa? Or is there a (set of) variable(s)  $z$  that drive(s) both? And can correlational data ever be used for causal inference? These questions were addressed in a relatively recent article in this journal (Lee, 2012) along with excellent open peer commentary. Here, I argue that the plausibility of any causal interpretation of personality trait-outcome associations hinges not just on directionality of influence or effect size but also on the plausibility of the underlying idea that traits exist in a holistic form that has causal potential.<sup>1</sup>

#### *The basis for causal interpretation in personality research*

For a phenomenon to exert some kind of causal force, I argue, it has to have some kind of existential reality and coherence. Among other things, this means that particular broad personality

---

<sup>1</sup> This article is concerned with causal explanations of trait-outcome associations and not merely prediction of outcomes from personality traits. Explanation is not necessarily the same thing as prediction. For example, one can predict that countries with high self-reported Conscientiousness have low levels of national wealth and low life expectancy (Möttus, Allik, & Realo, 2010), but this does not necessarily mean that high average Conscientiousness is involved in making the average person poor or shortening his or her life. In some settings (e.g., organisational or clinical contexts), though, mere ability to predict an outcome from personality variables has importance.

traits such as Extraversion and Conscientiousness have to exist independently of how researchers have happened to conceive and measure them to warrant being granted causal powers of their own. Low Conscientiousness as such can have causal effects on mortality rates only if it is a real human attribute that can, at least in some conditions or in some people, initiate the kinds of behaviors or biological processes that may eventually shorten life. When I use the term ‘low Conscientiousness as such’, or any other trait, I refer to it as a real psychobiological attribute rather than a shorthand summary of various behaviors, thoughts, feelings and whatever else that happen to correlate and we have decided to use for operationalizing personality. This interpretation of traits corresponds to what has been called 'realist ontology' (Borsboom, Mellenbergh, & van Heerden, 2003, p. 203).

Of course, our personality operationalizations and their associations with outcomes are ever only models of reality and models are always approximations and abstractions—even when what they model is real. But models differ in the levels of concreteness with which they are intended to describe tangible reality. Broad personality traits such as those of the FFM are often treated as psychological and biological realities that are “in our skins”, quite unlike how one might think of some otherwise equally legitimate abstract concepts such as moral consequences of the French Revolution. This is exemplified by interpretations of the broad personality traits as “affecting”, “leading to” or being “determinants of” measurable behavior and thereby concrete life outcomes, on the basis of observed associations in particular samples of people, as well as by attempts to establish their biological basis in people’s brains (e. g., Bjørnebekk et al., 2013; DeYoung, 2015) and genomes (e. g., de Moor et al., 2012). The very existence of such research is premised on the belief that particular traits exist within individuals' brains. Few would scan people's genomes and brains to look for moral consequences of the French Revolution.

This realist assumption about traits is most naturally consistent with the common-cause latent trait interpretation (Borsboom et al., 2003; Markus & Borsboom, 2013). In the psychometric framework, such traits are also called ‘reflective’ (Edwards & Bagozzi, 2000). According to this



view, traits can be thought of as unobserved 'generators' (underlying common causes) that, perhaps in transaction with situational circumstances, produce the observable regularities in behaviors, thoughts and feelings that coalesce as traits. Although the generators may be unobserved and can only be inferred from their manifestations (indicators), they exist independently of their observable indicators and cause them. It is thus the individual differences in the powers of these generators that matter for both individual differences in the observable trait manifestations and the outcomes to which the traits are linked: indicators are associated with the outcomes only because they all share the same causal input from the underlying generators. A less powerful generator means less of both the manifestations and any of the outcomes to which it contributes. John's Extraversion (generator) is set at a lower rate than Jane's and therefore he is less likely to attend parties and tends to be less talkative in many social settings (manifestations), and if he also has fewer one-night stands (outcome) it is possibly his relatively slow generator that contributes to this and not specifically his paucity of party attendance or tendency to reticence<sup>2</sup>. Such traits are holistic in the sense of not being reducible to their single manifestations or even their composites.

*Realist common-cause interpretation of traits is common*

Although not all researchers may, at least explicitly, commit to this full exposition of the generator interpretation of traits, the general idea appears to be consistent with the default trait conceptualization. For example, this view is in line with the influential Five-Factor Theory, which conceives traits as “biologically based properties of the individual that affect the rest of the personality system, but are not themselves affected by it” (McCrae & Costa, 2008, p. 278). Likewise, according to the Cybernetic Big Five Theory (DeYoung, 2015) a personality trait corresponds to “the typical functional level of the underlying psychological processes responsible for generating the emotional, motivational, cognitive, and behavioral states associated with that

---

<sup>2</sup> This example illustrates the difficulties related to distinguishing outcomes from the kinds of indicators that we often use for operationalizing our purported traits: we could also think of the number of one-night stands as a trait-indicator and attending parties as an outcome.

trait” (p. 37). Roberts and colleagues (2014) have proposed that Conscientiousness “represents a coherent psychobiological construct that influences the multiple facet manifestations” (p. 1318). Guillaume-Hanes, Morese and Funder (2012) have said that careful driving and avoiding binge drinking are related because they “manifest the same underlying latent trait of Conscientiousness” (p. 439).

For now, I do not challenge this realist interpretation. For the best part of this article, I am assuming that we have sufficiently good reasons to believe that, in principle, personality traits as such exist and can exert forces outside the personality domain in real and holistic manners. The following sections address some of the implications of this assumption.

*What is necessary to have existential and holistic reality and thus causal power?*

When we hypothesize there to be genes, brain correlates or behavioral consequences of broadly defined traits, say, Agreeableness or Openness, I maintain that we at least implicitly assume the traits to be unitary in both their etiology and any causal forces they exert on other phenomena. That is, the ostensible manifestations used to define a purported trait (e.g., its items or facets) have to have overlapping causes (e.g., genes or brain correlates) and consequences (e.g., outcome correlations), to the extent that they reflect this trait. It would not make much sense to relate scores of a purported trait to genes, brain parameters and life outcomes if the trait was actually a collection of behaviors, thoughts and feelings that have different causes and exert different kinds of forces. If this were the case, it would be more useful to look for the genes, brain correlates or outcome correlations of these specific characteristics. In other words, trait scores should be linked with their causal factors or consequences in similar ways regardless of the specific observable indicators that are used to operationalize them in any given study. Otherwise, it is something specific to trait operationalizations that matters, rather than the trait themselves. This corresponds to what Spearman (1927) called *the theorem of indifference of indicator*. Also, the manifestations of etiologically unitary traits should display similar developmental trajectories and within-individual

fluctuations, and should respond similarly to experimental manipulations, to the extents that they reflect the traits (Cattell, 1946). Speaking of Cattell, for him it was exactly the traits with evidence for the highest degree of unity that could demonstrate the greatest “usefulness and efficiency in all kinds of explanatory and predictive situations” (Cattell, 1946, p. 101).

This requirement of etiological and causal unity is consistent with the often-used common-cause latent variable interpretation of traits (Borsboom et al., 2003; Markus & Borsboom, 2013). Under this interpretation, unidimensional traits (those reflecting single common causes) exist independently of their indicators such as items or facets, so the indicators are exchangeable—at least in principle (Bollen & Lennox, 1991). This justifies that researchers use different questionnaires and yet claim to have measured the same FFM traits (Pace & Brannick, 2010). Individual differences in Extraversion (the generator), for example, could be inferred from individual differences either in how talkative and fun-seeking or how assertive and socially outgoing people are. No single manifestation is a perfect indicator of its ostensible parent trait because the trait is more broadly defined, and any specific manifestation may reflect some other sources of substantive variance, in addition to random error. Hence, traits are typically operationalized as the common variances of their multiple indicators, assuming that it is these common variances that best reflect individual differences in the traits as such. To the extents, then, that the indicators largely reflect their ostensible parent traits and that it is indeed the traits as defined by these common variances that have particular etiological factors or exert causal forces on outcomes, their indicators should display largely similar associations with these etiological factors and outcomes—for trait-pertinent causal interpretations to be feasible.

Of course, not all indicators of a trait need to be correlated with the outcome at hand to *exactly* the same degree. First, indicators can differ in the degree to which they reflect their presumed latent parent trait (factor loadings) and the associations between single indicators and outcomes should therefore be proportional to factor loadings. Second, indicators may contain

specific variance that is correlated with the outcome over and above the variance they share with other indicators of the same purported trait. However, to say that it is the trait as such that is correlated with the outcome, it seems necessary that most individual indicators have at least some above-chance association with the outcome in the same direction. The same logic, of course, applies to correlations with factors hypothesized to cause variation in the trait itself.

*Holistic collections of behavioral patterns could exist without common causes*

In principle, of course, one might argue that a collection of behavioral patterns that we have labeled a trait, say Neuroticism, consists of a fixed set of irreplaceable constituents, or at least a finite list of core constituents. In this case, there would be no inevitable need for a *single* common cause for the inter-correlated manifestations of the trait. For example, McCrae (2015) proposed an alternative to the common-cause interpretation of traits by suggesting that traits could be seen as unions of their semi-autonomous constituents. In this view, the inter-correlations of trait constituents would result from only partly overlapping sets of etiological factors (no etiological factor would then pertain to all constituents but only a subset of them, but the subsets would partly overlap), and the unique etiological factors would make the constituents partly autonomous. Or, a trait could be holistic because its (core) constituents have evolved to work in concert to allow the organism to achieve some survival or reproductive purpose. As is the case with common-cause latent traits, such a collection of behavioral patterns would not be reducible to any of its single constituents nor even their composite. For example, one could draw a parallel with football teams, which consist of exactly eleven players with different more or less clearly defined roles, or with aircraft, which have a fairly standard set of core features (fuselage, wings, tail, engines etc.) as well as numerous optional characteristics. For a football team to be successful, the skills of individual players matter, but it is the players' co-ordinated actions that may matter at least as much or even more. Likewise, for an aircraft to fly as intended, its core features have to perform well individually, but they also have to work in concert.

Would it be possible to say that Neuroticism is defined by a fixed and exhaustive set of  $k$  (core) items, no more and no less? I am not aware of such claims having been made for personality traits. McCrae (2015), who discussed traits as unions of only partly overlapping constituents, also appeared cautious on this question. Could it be that Neuroticism is a union of particular constituents that have to work in concert, giving the trait emergent properties? I am not aware of such claims either. But even if such claims were made and appeared plausible, the requirement of causal unity would probably stand, regardless of whether traits are unions due to partly overlapping etiological factors of their constituents or because they form some kind of functional unities with emergent properties. If only one bit of the union is linked to the outcome, causality pertains to this bit rather than the whole, even if this small bit is crucial for the functioning of the whole.

Hypothetically, suppose that there was evidence that a surge in the failure of a particular core feature of aircraft, say landing gear, was linked to the majority of increasingly frequent crash-landings, regardless of other aircraft characteristics, including their general manufacturers who tend to outsource landing gears to independent parts manufacturers. In such instances, one could naturally ascribe a relatively higher proportion of causality for the increased accident rate to some defect in some aircraft landing gears, as opposed to the aircraft as such or their general manufacturers. It would probably not make sense to conclude that some aircraft as such have become unsafe overall and need to be redesigned completely, because variability in most of the aircraft parts is irrelevant. One could blame aircraft design or those of particular types of aircraft as such if the sources of accidents were broader and/or spread across the aircraft features, pointing to general assembly problems, for example. Instead, attempts to reinstate flight safety should focus on the particular suppliers and/or characteristics of the unsafe landing gears, because some change involving them caused the surge of accidents. Also hypothetically, if there was evidence that investment in a particular player position on football teams, say, goalkeeper (supposing investment reflects players' quality), is correlated with the number of wins teams have, one would naturally

consider ascribing a relatively higher proportion of causality for game wins to goalies than to other types of players or teams in general. Managers would then be well advised to invest in goalies in particular to increase their teams' success rates.

*What does this mean for research programs focused on identifying personality effects on outcomes?*

It is widely acknowledged that broad traits such as those of the FFM are comprised of at least somewhat distinguishable components such as aspects (DeYoung, Quilty, & Peterson, 2007) or facets (Costa & McCrae, 1992) that reflect *additional* aspects of personality beyond the extents to which they reflect their respective broad traits (usually interpreted as common causes in these conceptualizations). And indeed, aspects and (especially) facets are often employed when personality-outcome correlations (Roberts, Chernyshenko, Stark, & Goldberg, 2005) or age-differences (Terracciano, McCrae, Brant, & Costa, 2005) in personality traits are investigated, although little to no attention is typically paid to the associations at the level of the items that make up the traits, aspects and facets. But even if aspect- or facet-level associations are reported, the theoretical implications of such analyses for causal interpretations are rarely discussed (one attempt to do this can be found in Judge, Rodell, Klinger, Simon, & Crawford, 2013). For example, suppose that only a subset of facets of an FFM trait is linked to an outcome, but this subset also makes the FFM trait—the sum score to which these facets contribute—appear correlated with it. This is not a rare occurrence in existent literature, as will be noted below. In such a case, I maintain, there is unlikely to be sufficient evidence to conclude that it is the FFM trait as such that is linked to the outcome, and it is even less likely that it is the FFM trait that has some causal impact on the outcome.

I argue that whenever a trait, whether measured as a latent variable or a sum score, is linked to an outcome, consistency of the associations among the items and, where available, facets of the trait-operationalization should be tested. Moreover, I argue that the results of these analyses should be interpreted in light of their theoretical implications for the nature of the observed associations. If

the associations are consistent across items and facets, they are more likely to pertain to the trait: this a necessary, although not a sufficient, precondition for causal interpretation (e.g., this does not say anything about the directionality of the association). In contrast, if the associations are specific to only some items or facets, then they are unlikely to pertain to the trait and would be more appropriately interpreted as pertaining to the (unique) variance of the specific items or facets involved. Such sensitivity analyses should not be considered bonuses to trait-level analyses, but as essential requirements for making the trait-level analyses interpretable. Simply reporting FFM trait-outcome correlations is not sufficient, in my view.

Such sensitivity analyses should also be carried out when trait scores are linked with potential causal sources such as genetic variants, brain parameters, or demographic variables. For example, we systematically compared the developmental trajectories of the facets of the same FFM traits and items of the same FFM facets and found that facets of all traits and items of all facets varied significantly in how they were correlated with age (Möttus et al., 2015). This pointed to the possibility that different causal forces acted upon different aspects of the same traits and facets to varying degrees throughout the lifespan, calling their coherence and existential realism into question.

#### *Samples of people and samples of stimuli*

My argument for causal unity of traits in relation to outcomes (or etiological unity in relation to the possible causes of trait-variation, for that matter) is consistent with a recent discussion of interpretation problems in experimental studies (Westfall, Judd, & Kenny, 2015; Judd, Westfall, & Kenny, 2012). In particular, these authors noted that experimental research often fails to take into account stimulus variation—that the stimuli used in studies are never perfect exemplars or operationalizations of the underlying constructs of interest. In many cases, stimuli (e.g., photographs shown or words presented in an experiment) should be considered potentially interchangeable, drawn from theoretically appropriate populations of relevant stimuli—exactly as

we think of participants. Failure to consider and empirically address this requirement may lead to incorrect conclusions in situations where observed associations are specific to particular stimuli chosen to operationalize the phenomena of interest. To prevent this, Westfall and colleagues (Westfall, Judd, & Kenny, 2015; Judd, Westfall, & Kenny, 2012) proposed treating experimental stimuli as random rather than fixed factors and employing appropriate mixed-effect modeling approaches.

Analogously, we can think of personality test items and perhaps facets as stimulus material. Just as analogously, then, we should treat them as random and interchangeable samples from the corresponding universes of theoretically appropriate stimuli. For example, given some existentially and holistically real latent Neuroticism trait, for example, the 48 items of the Neuroticism scale of the Revised NEO Personality Inventory (NEO-PI; Costa & McCrae, 1992) are probably not the only possible 48 indicators of Neuroticism as it occurs in nature: they only comprise a sample of items measuring this ostensible underlying trait. Indeed, other FFM scales include different items for measuring the same construct (Goldberg, 1999). No sample of indicators perfectly reflects the intended trait as such, but we should do our best to ensure as good an approximation as possible and a viable way to do this is to compare results based on different samples of indicators.

### *Existing evidence*

Because researchers rarely report item-level correlations, it is difficult to estimate to what degree trait-outcome associations tend to be consistent across the items used for operationalizing traits. One exception is the study of Terracciano and colleagues (2009), who reported that although the Impulsiveness facet of the NEO-PI had the strongest link with high body mass index (BMI), the association was driven by only the two items of the facet scale that specifically referred to over-eating. That is, had the NEO-PI constructors not chosen to include two eating items in the Impulsiveness scale, its scores would probably not have been associated with being overweight. Although Terracciano and colleagues “adopted the perspective that personality traits influence



anthropometric values” (p. 687), it seems only too likely that at least for Impulsiveness it was not the broadly construed construct as such but more specifically tending to eat too much that was the correlate and thereby possibly a causal factor of high BMI. Similar findings were reported by Vainik and colleagues (2015). However, other studies that have reported associations between the Impulsiveness scale and BMI (e.g., Sutin, Ferrucci, Zonderman, & Terracciano, 2011; Sutin et al., 2013) have failed to address item-level correlations, which renders interpretations of their findings ambiguous, given the results of these other studies.

Similarly, Gale and colleagues (submitted for publication) reported cross-sectional associations between the FFM traits and frailty in a large sample of Britons aged between 60 and 90 years. Conceptually, frailty reflects decreases in resistance to stressors and ability to maintain homeostasis in older age. In this study, it was indexed as a composite of 44 items measuring various aspects of health, as well as sensory and cognitive functioning. Lower Extraversion and Conscientiousness were associated with higher frailty scores, but the association with Extraversion was entirely driven by two items referring to being 'active' and 'lively', and the association with Conscientiousness pertained to a large extent (but not completely) to an item referring to 'hard-working'. In this case, it was not particularly meaningful to conclude that frailty was linked with Extraversion as such—it was more likely linked with the variance in being active and lively that was not shared with the other indicators of Extraversion.

Of course, in many—perhaps even most—instances the outcome associations do generalize across items of their respective scales, making it more likely that they pertain to the shared variance of trait indicators and thereby to the traits as such. But this needs to be demonstrated and not assumed. For example, Vainik and colleagues (2015) found that the associations of the Assertiveness and Order facets of the NEO-PI with BMI generalized across all items of the scales. Perhaps importantly, however, these are facets rather than any of the broad FFM traits, and they are not even considered facets of the *same* broad traits. This suggests that the associations could not be

generalized to the two relevant FFM traits.

Relative consistency of facets-within-trait associations with outcomes can also be evaluated, if facets are measured. To the extent the FFM traits have existential and holistic realism in the sense of causing rather than summarizing the observable and inter-correlated behaviors, thoughts and feelings, they should also exist independently of which particular set of facets researchers have carved out for them. To the best of my knowledge, no one has made a case that the six, say, Extraversion facets of the NEO-PI exhaust the theoretical population of facets for this trait. But even more importantly, if FFM traits were sums rather than causes of their constituent facets (i.e., formative traits; Edwards & Bagozzi, 2000), then it is the latter that would take causal priority. Given this, save for variability in the degree to which facets reflect their parent FFM traits (i.e., factor loadings), they should have similar correlations with outcomes, if it is the FFM traits as such and not the unique variance in facets that are linked with the outcome at hand. The incremental value of facets must lie in their unique variance on top of common trait-variance rather than in facets defining identifiably separable components of the traits.

There is considerable evidence that different facets of the FFM traits can have quite different correlations with outcomes, sometimes even in opposite directions. Because the factor loadings of facets are not typically reported, it is difficult to judge to what degree differences in how the facets reflect their parent FFM traits can account for this variability, though presumably this would never account for associations running in different directions. As with items, BMI has shown different associations with the facets of particular FFM traits. For example, although high BMI was correlated with high Neuroticism in Sutin and colleagues' (2011) study, this association was driven by only the Impulsiveness facet of the trait; and based on the above-discussed evidence from other studies, even this association could have been specific to the two eating-related items of the scale, though this was not addressed there. Likewise, the facets of Conscientiousness differed in their links with BMI, to the extent that they were in different directions (Sutin et al., 2013, 2011). Observations

that different facets of the same FFM trait had very different and sometimes even opposing outcome associations or that the apparent associations were driven by only a few facets of the trait have also been reported for other outcomes. Some relevant examples include job-performance (Judge et al., 2013; Tett, Steele, & Beauregard, 2003), antisocial behavior (Corff & Toupin, 2009), dietary habits (Mõttus, Realo, Allik, Deary, et al., 2012), smoking cessation (Fernández del Río et al., 2015), consumer behavior (Otero-López & Villardefrancos Pol, 2013), sexually transmitted disease (Mõttus, Realo, Allik, Esko, & Metspalu, 2012), relationship status and quality (Noftle & Shaver, 2006), and life-satisfaction (Marrero Quevedo & Carballeira Abella, 2011). It may well be that in many, if not most, instances the variability in outcome associations is due to variability in factor loadings, but this needs to be tested.

*How to investigate variability in indicator-outcome correlations?*

The easiest way to investigate variability in how indicators (items or facets) of a trait are linked to an outcome is to eyeball item- or facet-level correlations. If only a subset of items or facets of a scale are linked to the outcome at hand, and especially if the associations run in different directions, it is not likely that the trait as such is involved in the association.

Another rule of thumb is that none of the items or facets should have a stronger correlation with the outcome than their parent trait, at least when the association is driven primarily by the trait. This is because, theoretically, items and facets are only imperfect indicators of their parent traits. For example, if a latent Conscientiousness variable correlates .25 with variability in income and the item “I am self-disciplined” has a loading of .50 on the latent trait, then the item should have a correlation of at most  $.25 * .50 = .125$  with income. If, however, this single item has a notably higher correlation with income, then it is possible or even likely that the trait-outcome correlation of .25 is largely driven or at least notably inflated by this single, supposed-to-be-exchangeable item.

To the extent that there is concern that a trait-outcome association may be driven or inflated by specific individual indicators of the trait, it would be natural to drop these indicators and

recalculate the association. If the association remains similar, there is no reason to be further concerned. If the association is remarkably diminished, but still observable beyond what is likely due to chance, it should be interpreted as having been inflated by the specific trait indicators. If the association vanishes altogether, then it most likely pertained to these specific indicators and not to the trait itself.

Of course, dropping indicators from trait operationalizations may entail reductions of reliability, which also tends to weaken observable trait-outcome associations. It is therefore advisable to replace the dropped indicators with alternative indicators. One way to do this is to start with larger samples of indicators than is strictly necessary for sufficiently reliable measurement of the trait and draw different samples of indicators from these larger indicator populations, aggregate them and calculate their associations with outcomes. If a correlation pertains to the underlying trait, it should remain relatively similar across different draws of indicators. It is often possible to apply this approach to already-collected data. For example, the NEO-PI measures each FFM trait with 48 items—probably more than is necessary for reliably measuring a trait. It is therefore possible to draw samples of, say, 10 to 20 items from among the 48 and calculate the correlations of the resulting scales with the outcomes at hand. Mõttus and colleagues (2015) applied this approach to investigate the associations of FFM traits with age and found that the correlations varied substantially across scales comprising different draws of items of the same traits. Another way to sample different items for measuring the same underlying trait is to administer in parallel two questionnaires purportedly measuring this trait—e.g., the NEO-PI and IPIP (Goldberg, 1999)—and compare trait-outcome-correlations across the questionnaires.

Admittedly, there are too few appropriate formal methods for testing whether the associations of traits with outcomes are independent of which indicators are employed rather than being specific to particular indicators. Vainik and colleagues (2015) developed a formal procedure based on the idea of systematically dropping trait indicators and recalculating the associations. Another formal

method for testing the causal unity of traits is to use random effects models and treat indicators as random factors, as suggested by Judd and colleagues (2012). However, neither of these methods may efficiently account for variability in indicators' factor loadings. Therefore, attempts to develop more refined formal tests of causal unity of traits should be strongly encouraged.

### *Risk of triviality*

In instances where observed trait-outcome associations appear specific to particular indicators, I claim that they should *not* be interpreted as pertaining to the traits as such. Instead, the associations should be interpreted as pertaining to the specific constituents of the trait-operationalizations (indicators) that display the associations. However, sometimes these associations may be so specific as to be trivial or tautological. For example, if we find out that only those items of Neuroticism in general, and its Impulsiveness facet in particular, that directly refer to eating too much correlate with being overweight, this may not do much to improve our understanding of either personality or obesity. Likewise, finding out that the associations of Extraversion with well-being, life-satisfaction and happiness are to notable extents driven by the items of the Extraversion scale that directly refer to experiencing positive emotions (Marrero Quevedo & Carballeira Abella, 2011; Schimmack, Oishi, Furr, & Funder, 2004) may seem somewhat tautological. But so be it. There is no point in generalizing associations to levels of the trait-hierarchy to which they do not appear to apply, given the evidence at hand.

### **What if traits do not have existential realism to start with?**

The requirement that trait-outcome associations be independent of how the trait happens to be operationalized in any given study only applies when we assume that traits should take some real and holistic form. Perhaps some researchers do not assume that personality traits such those of the FFM are real psychobiological attributes that either exist independently of their operationalizations or are some kind of natural unions of their exhaustive sets of (core) features. Perhaps they interpret trait scores as aggregates of researcher-chosen constituents (e.g., Cramer et al., 2012; Wood,

Gardner, & Harms, 2015). For example, traits may be otherwise artificial constructions that serve only our practical research or assessment purposes such as summarizing individual differences succinctly (for a related discussion with respect to psychiatric disorders, see Kendler, Zachar, and Craver, 2011). If so, attributing causality to traits as such seems equally questionable regardless of whether their constituents have similar or different associations with the outcome at hand. Even if the associations generalize across trait constituents, causal interpretations may be more fruitfully based on these constituents rather than the summary-level traits.

To see why, it might be instructive to draw a parallel with socioeconomic status (SES), which could be defined as a composite score of educational level, income, occupational status and the quality of one's residence, for example. All of these constituents of SES tend to be inter-correlated and this composite, as well as similar others, have many well-known associations with important life outcomes (e.g., Mackenbach et al., 2008). Yet it would be presumptuous to *assume* that this is because these constituents come together from some common set of causes to create some kind of existentially and holistically real entity (the generator) that can in turn exert causal forces on all the associates of SES. There may be no existentially real SES-attribute on which people differ and which causes them to graduate from good schools, obtain prestigious jobs, earn high salaries and live in nice neighborhoods full of large, expensive homes.

I do realize that some may propose, say, intelligence or perhaps some parental characteristic as the common cause of SES indicators. In my view, however, it is at least equally plausible that the situation is more complex than that. I suspect that a good education, which is indeed predictable from pre-existing intelligence as well as parental education, helps one to obtain a prestigious job, most of which also bring rather high income and thereby ability to purchase a comfortable home in a good neighborhood, and most people indeed put reasonably high proportions of their financial resources into their homes, as well as tending to prefer to live among neighbors rather similar to themselves in income and education. Yes, high intelligence may provide extra help at every stage,

but it may also be itself boosted and/or maintained by the characteristics used to define the attained SES such as educational and occupational levels (Deary & Johnson, 2010; Ritchie, Bates, Der, Starr, & Deary, 2013). But it may sometimes also be that, despite their inter-correlation, each of these constituents substantively adds its own independent variance to what we consider SES. For example, sports champions and movie stars often earn very high incomes and have high social prestige, but may not be particularly well educated. And extremely highly-educated university professors do not tend to make anywhere near as much money as sports champions and movie stars, or even American football coaches at the same universities. Now, if SES actually coalesces due to causal links among its constituents or is a mere sum of these, it is not straightforward to interpret *it* as a cause of its associated outcomes—even if all the constituents of the SES operationalizations track with them in similar ways. This is because SES itself is then the *result* rather than the cause of its constituents and, unless it takes on emergent properties, it thereby owes its outcome correlations to these constituents.

If SES itself is not directly causal for either of these reasons, why, then, might low SES be associated with diabetes, for example? It seems possible that some of its constituents such as, perhaps most prominently, low educational attainment and income, along with whatever upstream antecedents these may have (e.g., intelligence; Möttus, Luciano, Starr, & Deary, 2013), can contribute to diabetes. This may happen, for example, via poor uptake of knowledge about health management, treatment, and early disease symptoms; greater exposure to stressful conditions; and limited access to resources that would help to sustain healthy lifestyle and effectively manage the precursors and symptoms of the condition.<sup>3</sup>

---

<sup>3</sup> Lutfey and Freese (2005) discuss SES as being the 'fundamental cause' of health outcomes such as diabetes. In their treatment, SES-outcome associations refer to large and persistent collections of pathways that link the constituents of SES to a variety of aspects of the outcomes (e.g., aspects of diabetes etiology and management). It is exactly the multiplicity and thereby persistence of the pathways that makes the role of SES in outcomes fundamental: such associations are extremely hard to break by knocking out single or a few pathways between SES and outcome constituents—there are simply too many of them and they tend to reproduce. It is not clear, however, to which extent one really needs to evoke SES as such for the causal explanations. It is possible that many of these multiple pathways that Lutfey and Freese (2005) discuss as linking SES to its associated outcomes are reducible to some particular

If one also interprets personality traits such as Conscientiousness as collections of etiologically and functionally autonomous behaviors, thoughts and feelings—albeit still often inter-correlated to varying degrees<sup>4</sup>, perhaps due to direct causal inter-relations with each other (Cramer et al., 2012) or due to shared functional values (Wood et al., 2015)—then their associations with outcomes could arise in analogous ways. Perhaps it is the specific constituents of our operationalizations of Conscientiousness that individually and thus cumulatively contribute to the kinds processes that help some people live longer. Having the habit of carefully checking for approaching cars before crossing the road and not exceeding the speed limit (possible constituents of NEO-PI Deliberation and Dutifulness facets) may decrease the likelihood of being involved in a traffic accident, whereas having enough self-discipline and organization skills (possible constituents of NEO-PI Self-Discipline and Order facets) for continuously exercising and monitoring diet may help maintain better metabolic state. These behaviors are likely to be inter-correlated, which may at least in part be due to serving the same purpose—a decision to take care of one's health and safety, for example, which would then be more appropriate to consider an underlying 'cause' of longevity. Yet making the bed in the morning or putting dirty clothes in a hamper may not be directly related to longevity, let alone cause it, even if they correlate with exercising and dieting (and thereby indirectly with road-checking and sensible driving), because they all contribute towards the common purpose of pleasing one's partner, for example. For a thoughtful treatment of these questions, refer to Wood and colleagues (2015).

---

constituents of SES such as, most likely, education and income (and their more distal precursors such as intelligence, for example). If so, SES itself may again be little more than a shorthand summary of numerous causal agents rather than the cause itself. Importantly, the SES-outcome associations would still need to be similarly observable across most of the pathways linking the constituents of both (Lutfey & Freese, 2005). Also importantly, it may be possible to break the SES-health associations after all: there is some evidence that the SES-mental health links can be persistently weakened in children by quasi-experimentally increasing just one of the SES constituents—family income (Costello, Erkanli, Copeland, & Angold, 2010).

<sup>4</sup> Note that the correlations among the items of the same FFM domains are typically relatively modest, with average correlations ranging somewhere from .10 to .20 or slightly higher. With 48 items, these average inter-item correlations result in impressive internal consistency estimates (Cronbach's alphas) of .84 and .92 or higher, however.



### **Some further considerations**

#### *Traits need to be defined independently of their mediators*

Simply documenting an association between a broad personality trait and an outcome may not be sufficient. Findings that Agreeableness correlates negatively with having a criminal record (Mõttus, Guljajev, Allik, Laidra, & Pullmann, 2012) or Conscientiousness correlates with brain integrity (Booth et al., 2014) have to be furnished with more specific mechanisms, either if these associations are to be used for attempts to reduce antisocial behavior or improve brain health, or for understanding how exactly personality exerts its influence on other life domains. It may be tempting to work out the mediating mechanisms by identifying the specific contents of personality scales that correlate with the outcomes particularly highly. For example, one may hypothesize that low NEO-PI Conscientiousness predicts high BMI because people low on this trait have limited self-discipline and are not well organized (Sutin et al., 2011), and thus they manifest this by not keeping good track of what they eat, scheduling in time for exercise, or monitoring their weight. On one hand, this reasoning seems to make perfect sense: low Conscientiousness is causal to its manifestations such as low self-discipline and lack of orderliness, which in turn cause some proximal factors contributing to being overweight. On the other hand, though, this reasoning runs the risk of being tautological. This is because self-discipline and orderliness are not logically independent from Conscientiousness: the trait has been in part operationalized using these proposed mediators. It is only when the potential mediators are dropped from the trait operationalizations that the proposed mediation pathways become more plausible. However, it seems all too likely that when these mediator-candidates are dropped from trait operationalizations, the trait-outcome links are substantially weakened or may sometimes vanish altogether.

#### *Traits should exist independently of outcomes*

A requirement for a good explanation is that the phenomenon being explained (explanandum)

be conceptually distinct from the phenomena that are used for the explanation (explanans). Therefore, outcomes have to be conceptually distinct from traits. This requirement may not always be met in studies examining trait-outcome associations.

Sometimes, the conceptual overlap between traits and outcomes is obvious. For example, consider personality traits such as Neuroticism and Extraversion on one hand and happiness and well-being on the other. Feeling positive emotions is part of how Extraversion is defined and measured in the FFM (e.g., the Positive Emotions facet of Extraversion in the NEO-PI; Costa & McCrae, 1992) and lack of positive feelings is also an integral part of being unhappy. It is therefore not surprising that low scores on the Depression and high scores on the Positive Emotion facets are the most consistent correlates of well-being and happiness, among the facets of Neuroticism and Extraversion, respectively (Marrero Quevedo & Carballeira Abella, 2011; Schimmack et al., 2004). Or, consider the similarity between the items “I like most aspects of my personality” and “In general, I feel I am in charge of the situation in which I live” of the Scales of Psychological Well-Being (Ryff & Keyes, 1995) and, respectively, the items “I have a low opinion of myself” and “Sometimes things look pretty bleak and hopeless to me” of the Depression facet of the NEO-PI. Both of these measures were included in a meta-analysis concluding that there are both direct and indirect associations between personality traits with subjective well-being and that “long-term happiness is largely contingent on internal characteristics” (Steel, Schmidt, & Shultz, 2008, p. 151).

At other times, the overlap may be more subtle. In a sense, at least some personality ratings may have broad life outcomes such as educational and occupational success built right into them. Personality traits broadly summarize what people do and so do the outcomes. For example, getting ahead in academia or advancing one's career often require being hard-working, dutiful, orderly and deliberate across a range of circumstances as well as being perceived as such by self and others—the very same characteristics that Conscientiousness items are likely to reflect and/or people are likely to think about as they respond to these items. In some ways, a comprehensive personality test

is a little model of life as the respondent is perceived living it. Therefore, *some* overlap between broad personality trait scores and the life outcomes is almost guaranteed by definition. Of course, this does not apply when personality ratings are obtained before the outcomes could influence them.

Moreover, some more specific outcomes may be parts of how personality traits are defined or be the specific behaviors people think about when rating the traits. For example, consider the relation between dietary habits and Openness. Part of the definition of the trait is being willing to try new things, and trying food items that are not part of the typical local diet is one of the ways that people can try new things. In fact, a NEO-PI Openness item explicitly asks about diet (“I often try new and foreign foods”). As a result, it is perhaps rather unsurprising that there is evidence for a small correlation between Openness and adherence to less common types of diet (Möttus, Realo, Allik, Deary, et al., 2012). Likewise, physical activity is part of the definition of Extraversion (e.g., several of the NEO-PI Extraversion items address being active and having an active life-style, such as “I am a very active person”), which may contribute to the correlation between them (Rhodes & Smith, 2006).

Sometimes, it may be possible to reduce or completely remove direct trait-outcome overlap by skimming through personality test items and omitting ones that either clearly refer to the outcome at hand or that may inspire reflection on the outcome from statistical analyses. When this is not possible but there are still reasons to suspect some conceptual overlap between traits and outcomes, it may be advisable to factor this into interpretations of the associations. For example, it may be tautological to glorify broad personality traits because they correlate with some broad life outcomes when the trait operationalizations consist of behaviors directly required to achieve these outcomes, or when the trait ratings may in part reflect the very outcomes themselves. It would be surprising, one may argue, if summaries of what people have done (outcomes) did not correlate in some ways with summaries of what people say they typically do (personality test scores).

### *Implications for traits*

The extent to which broad traits display causal unity or fail to do so may have implications for their very nature. In particular, apparent causal unity would be consistent with traits having existential realism and being the kinds of generators posited by latent trait model. In contrast, widespread evidence that constituents of the operationalizations of broad traits show variability in links with outcome variables beyond variability in factor loadings would not be consistent with traits being causally and thereby existentially unitary, at least in the latent common-cause trait sense.

### **Conclusion**

My main arguments in this article were as follows. First, I argued that, for the purposes for which personality traits are often linked with outcome variables—to substantiate traits and explain outcomes, and especially when hoping to develop interventions to change the outcomes—the associations need to be capable of being interpreted in causal terms. Indeed, causal interpretations are widespread in the literature, although most researchers do acknowledge, often in cautionary notes restricted to study limitations sections, that their analyses do not fully justify these interpretations. Second, I argued that causal interpretations require traits to have existential realism and holism. Indeed, broad traits such as those of the FFM are often assumed to have such properties. Third, I argued that, for the causal interpretations to be possible, traits have to display evidence of causal unity: constituents of trait-operationalizations (or trait-indicators) have to be linked with outcomes in similar ways, save for variability in factor loadings. In other words, the associations should not depend on how traits happen to be operationalized. Fourth, I argued that such causal unity can be and should be tested in each and every study that seeks to link personality traits with outcomes, although formal methodology for doing this robustly requires further development. Existing evidence suggests that observed associations between, for example, FFM traits and outcomes at least sometimes depend on which particular constituents have been included

in trait-operationalizations, pointing to potential lack of causal unity of the FFM trait and thereby calling causal interpretations into question. However, this has rarely been explicitly considered in the literature so far. I thus call for greater theoretical and methodological rigor when linking personality traits with outcomes and provide one conceptual framework for doing so.

### **Acknowledgements**

I am grateful to Tom Booth and Aja Murray for their comments on an earlier version of this manuscript, and to Wendy Johnson for her help throughout the preparation of the manuscript. I am also grateful to William Revelle and an anonymous reviewer for their helpful suggestions as well as to Uku Vainik for persistently provoking me to think of the questions addressed in this article.

## References

- Ashton, M. C., & Lee, K. (2007). Empirical, theoretical, and practical advantages of the HEXACO model of personality structure. *Personality and Social Psychology Review*, *11*, 150–166. <http://doi.org/10.1177/1088868306294907>
- Bjørnebekk, A., Fjell, A. M., Walhovd, K. B., Grydeland, H., Torgersen, S., & Westlye, L. T. (2013). Neuronal correlates of the five factor model (FFM) of human personality: Multimodal imaging in a large healthy sample. *NeuroImage*, *65*, 194–208. <http://doi.org/10.1016/j.neuroimage.2012.10.009>
- Bollen, K., & Lennox, R. (1991). Conventional wisdom on measurement: A structural equation perspective. *Psychological Bulletin*, *110*, 305–314. <http://doi.org/10.1037/0033-2909.110.2.305>
- Booth, T., Möttus, R., Corley, J., Gow, A. J., Henderson, R. D., Maniega, S. M., ... Deary, I. J. (2014). Personality, health, and brain integrity: the Lothian birth cohort study 1936. *Health Psychology*, *33*, 1477–1486. <http://doi.org/10.1037/hea0000012>
- Borsboom, D., Mellenbergh, G. J., & van Heerden, J. (2003). The theoretical status of latent variables. *Psychological Review*, *110*, 203–219. <http://doi.org/10.1037/0033-295X.110.2.203>
- Caprara, G. V., Schwartz, S., Capanna, C., Vecchione, M., & Barbaranelli, C. (2006). Personality and politics: Values, traits, and political choice. *Political Psychology*, *27*, 1–28. <http://doi.org/10.1111/j.1467-9221.2006.00447.x>
- Carlo, G., Okun, M. A., Knight, G. P., & de Guzman, M. R. T. (2005). The interplay of traits and motives on volunteering: agreeableness, extraversion and prosocial value motivation. *Personality and Individual Differences*, *38*, 1293–1305. <http://doi.org/10.1016/j.paid.2004.08.012>
- Cattell, R. B. (1946). Personality structure and measurement. *British Journal of Psychology. General Section*, *36*, 88–103. <http://doi.org/10.1111/j.2044-8295.1946.tb01110.x>
- Corff, Y. L., & Toupin, J. (2009). Comparing persistent juvenile delinquents and normative peers with the Five-Factor Model of Personality. *Journal of Research in Personality*, *43*, 1105–1108. <http://doi.org/10.1016/j.jrp.2009.06.011>

- Costa, P. T., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual*. Odessa, FL: Psychological Assessment Resources.
- Costello, E. J., Erkanli, A., Copeland, W., & Angold, A. (2010). Association of family income supplements in adolescence with development of psychiatric and substance use disorders in adulthood among an American Indian population. *JAMA*, *303*, 1954–1960. <http://doi.org/10.1001/jama.2010.621>
- Cramer, A. O. J., van der Sluis, S., Noordhof, A., Wichers, M., Geschwind, N., Aggen, S. H., ... Borsboom, D. (2012). Dimensions of normal personality as networks in search of equilibrium: You can't like parties if you don't like people. *European Journal of Personality*, *26*, 414–431. <http://doi.org/10.1002/per.1866>
- Deary, I. J., & Johnson, W. (2010). Intelligence and education: causal perceptions drive analytic processes and therefore conclusions. *International Journal of Epidemiology*, *39*, 1362–1369. <http://doi.org/10.1093/ije/dyq072>
- Deary, I. J., Weiss, A., & Batty, G. D. (2010). Intelligence and personality as predictors of illness and death. *Psychological Science in the Public Interest*, *11*, 53 –79. <http://doi.org/10.1177/1529100610387081>
- de Moor, M. H. M., Costa, P. T., Terracciano, A., Krueger, R. F., de Geus, E. J. C., Toshiko, T., ... Boomsma, D. I. (2012). Meta-analysis of genome-wide association studies for personality. *Molecular Psychiatry*, *17*, 337–349. <http://doi.org/10.1038/mp.2010.128>
- DeYoung, C. G. (2015). Cybernetic Big Five Theory. *Journal of Research in Personality*, *56*, 33–58. <http://doi.org/10.1016/j.jrp.2014.07.004>
- DeYoung, C. G., Quilty, L. C., & Peterson, J. B. (2007). Between facets and domains: 10 aspects of the Big Five. *Journal of Personality and Social Psychology*, *93*, 880–896. <http://doi.org/10.1037/0022-3514.93.5.880>
- Edwards, J. R., & Bagozzi, R. P. (2000). On the nature and direction of relationships between constructs and measures. *Psychological Methods*, *5*, 155–174.
- Eysenck, H. J., & Eysenck, M. W. (1985). *Personality and individual differences: A natural science*

*approach*. New York: Plenum.

- Fernández del Río, E., López-Durán, A., Rodríguez-Cano, R., Martínez, Ú., Martínez-Vispo, C., & Becoña, E. (2015). Facets of the NEO-PI-R and smoking cessation. *Personality and Individual Differences, 80*, 41–45. <http://doi.org/10.1016/j.paid.2015.02.030>
- Lutfey, K., & Freese, J. (2005). Toward some fundamentals of fundamental causality: Socioeconomic status and health in the routine clinic visit for diabetes. *American Journal of Sociology, 110*, 1326–1372. <http://doi.org/10.1086/428914>
- Gale, C., Möttus, R., Deary, I. J., Cooper, C., & Sayer, A. A. (submitted for publication). Personality and risk of frailty: the English Longitudinal Study of Ageing. *Annals of Behavioral Medicine*.
- Goldberg, L. R. (1990). An alternative ‘description of personality’: The Big-Five factor structure. *Journal of Personality and Social Psychology, 59*, 1216–1229. <http://doi.org/10.1037/0022-3514.59.6.1216>
- 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. J. Deary, F. De Fruyt, & F. Ostendorf (Eds.), *Personality Psychology in Europe* (Vol. 7, pp. 7–28). Tilburg: Tilburg University Press.
- Guillaume-Hanes, E., Morse, P., & Funder, D. (2012). Network models in the organization of personality. *European Journal of Personality, 26*, 439–440. <http://doi.org/10.1002/per.1872>
- Hagger-Johnson, G., Bewick, B. M., Conner, M., O’Connor, D. B., & Shickle, D. (2011). Alcohol, conscientiousness and event- level condom use. *British Journal of Health Psychology, 16*, 828–845. <http://doi.org/10.1111/j.2044-8287.2011.02019.x>
- Jokela, M. (2009). Personality predicts migration within and between U.S. states. *Journal of Research in Personality, 43*, 79–83. <http://doi.org/10.1016/j.jrp.2008.09.005>
- Jokela, M., Batty, G. D., Nyberg, S. T., Virtanen, M., Nabi, H., Singh-Manoux, A., & Kivimäki, M. (2013). Personality and all-cause mortality: Individual-participant meta-analysis of 3,947 deaths in 76,150 adults. *American Journal of Epidemiology*, kwt170. <http://doi.org/10.1093/aje/kwt170>



- Jokela, M., Elovainio, M., Nyberg, S. T., Tabák, A. G., Hintsala, T., Batty, G. D., & Kivimäki, M. (2014). Personality and risk of diabetes in adults: pooled analysis of 5 cohort studies. *Health Psychology, 33*, 1618–1621. <http://doi.org/10.1037/hea0000003>
- Judd, C. M., Westfall, J., & Kenny, D. A. (2012). Treating stimuli as a random factor in social psychology: a new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology, 103*, 54–69. <http://doi.org/10.1037/a0028347>
- Judge, T. A., Rodell, J. B., Klinger, R. L., Simon, L. S., & Crawford, E. R. (2013). Hierarchical representations of the five-factor model of personality in predicting job performance: Integrating three organizing frameworks with two theoretical perspectives. *The Journal of Applied Psychology, 98*, 875–925. <http://doi.org/10.1037/a0033901>
- Kendler, K. S., Zachar, P., & Craver, C. (2011). What kinds of things are psychiatric disorders? *Psychological Medicine, 41*, 1143–1150. <http://doi.org/10.1017/S0033291710001844>
- Lee, J. J. (2012). Correlation and causation in the study of personality. *European Journal of Personality, 26*, 372–390. <http://doi.org/10.1002/per.1863>
- Mackenbach, J. P., Stirbu, I., Roskam, A.-J. R., Schaap, M. M., Menvielle, G., Leinsalu, M., & Kunst, A. E. (2008). Socioeconomic inequalities in health in 22 European countries. *The New England Journal of Medicine, 358*, 2468–2481. <http://doi.org/10.1056/NEJMsa0707519>
- Malouff, J. M., Thorsteinsson, E. B., Rooke, S. E., & Schutte, N. S. (2007). Alcohol involvement and the Five-Factor Model of personality: A meta-analysis. *Journal of Drug Education, 37*, 277–294. <http://doi.org/10.2190/DE.37.3.d>
- Malouff, J. M., Thorsteinsson, E. B., & Schutte, N. S. (2006). The Five-Factor Model of personality and smoking: A meta-analysis. *Journal of Drug Education, 36*, 47–58. <http://doi.org/10.2190/9EP8-17P8-EKG7-66AD>
- Markus, K. A., & Borsboom, D. (2013). *Frontiers of test validity theory: Measurement, causation, and meaning*. New York, NY US: Routledge/Taylor & Francis Group.
- Marrero Quevedo, R. J., & Carballeira Abella, M. (2011). Well-being and personality: Facet-level analyses. *Personality and Individual Differences, 50*, 206–211.

<http://doi.org/10.1016/j.paid.2010.09.030>

- McCrae, R. R. (2015). A more nuanced view of reliability: Specificity in the trait hierarchy. *Personality and Social Psychology Review*, *19*, 97–112. <http://doi.org/10.1177/1088868314541857>
- McCrae, R. R., & Costa, P. T. (2008). Empirical and theoretical status of the five-factor model of personality traits. In B. Boyle, G. Matthews, & D. Saklofske (Eds.), *The SAGE handbook of personality theory and assessment: Volume 1 — Personality theories and models* (pp. 273–295). London: SAGE.
- McCrae, R. R., & John, O. P. (1992). An introduction to the five-factor model and its applications. *Journal of Personality*, *60*, 175–215. <http://doi.org/10.1111/j.1467-6494.1992.tb00970.x>
- Mõttus, R., Allik, J., & Realo, A. (2010). An attempt to validate national mean scores of Conscientiousness: No necessarily paradoxical findings. *Journal of Research in Personality*, *44*, 630–640. <http://doi.org/10.1016/j.jrp.2010.08.005>
- Mõttus, R., Guljajev, J., Allik, J., Laidra, K., & Pullmann, H. (2012). Longitudinal associations of cognitive ability, personality traits and school grades with antisocial behaviour. *European Journal of Personality*, *26*, 56–62. <http://doi.org/10.1002/per.820>
- Mõttus, R., Luciano, M., Starr, J. M., & Deary, I. J. (2013). Diabetes and life-long cognitive ability. *Journal of Psychosomatic Research*, *75*, 275–278. <http://doi.org/10.1016/j.jpsychores.2013.06.032>
- Mõttus, R., McNeill, G., Jia, X., A, C., Starr, J. M., & Deary, I. J. (2013). The associations between personality, diet and body mass index in older people. *Health Psychology*, *32*, 353–360. <http://doi.org/10.1037/a0025537>
- Mõttus, R., Realo, A., Allik, J., Deary, I. J., Esko, T., & Metspalu, A. (2012). Personality traits and eating habits in a large sample of Estonians. *Health Psychology*, *31*, 806–814. <http://doi.org/10.1037/a0027041>
- Mõttus, R., Realo, A., Allik, J., Esko, T., & Metspalu, A. (2012). History of the diagnosis of a sexually transmitted disease is linked to normal variation in personality traits. *The Journal of Sexual Medicine*, *9*, 2861–7. <http://doi.org/10.1111/j.1743-6109.2012.02891.x>

- Mõttus, R., Realo, A., Allik, J., Esko, T., Metspalu, A., & Johnson, W. (2015). Within-trait heterogeneity in age group differences in personality domains and facets: Implications for the development and coherence of personality traits. *PLoS ONE*, *10*, e0119667. <http://doi.org/10.1371/journal.pone.0119667>
- Noftle, E. E., & Shaver, P. R. (2006). Attachment dimensions and the big five personality traits: Associations and comparative ability to predict relationship quality. *Journal of Research in Personality*, *40*, 179–208. <http://doi.org/10.1016/j.jrp.2004.11.003>
- Otero-López, J. M., & Villardefrancos Pol, E. (2013). Compulsive buying and the Five Factor Model of personality: A facet analysis. *Personality and Individual Differences*, *55*, 585–590. <http://doi.org/10.1016/j.paid.2013.05.005>
- Ozer, D. J., & Benet-Martínez, V. (2006). Personality and the prediction of consequential outcomes. *Annual Review of Psychology*, *57*, 401–421. <http://doi.org/10.1146/annurev.psych.57.102904.190127>
- Pace, V. L., & Brannick, M. T. (2010). How similar are personality scales of the ‘same’ construct? A meta-analytic investigation. *Personality and Individual Differences*, *49*, 669–676. <http://doi.org/10.1016/j.paid.2010.06.014>
- Rhodes, R. E., & Smith, N. E. I. (2006). Personality correlates of physical activity: a review and meta-analysis. *British Journal of Sports Medicine*, *40*, 958–965. <http://doi.org/10.1136/bjism.2006.028860>
- Ritchie, S. J., Bates, T. C., Der, G., Starr, J. M., & Deary, I. J. (2013). Education is associated with higher later life IQ scores, but not with faster cognitive processing speed. *Psychology and Aging*, *28*, 515–521. <http://doi.org/10.1037/a0030820>
- Roberts, B. W., Chernyshenko, O. S., Stark, S., & Goldberg, L. R. (2005). The structure of Conscientiousness: An empirical investigation based on seven major personality questionnaires. *Personnel Psychology*, *58*, 103–139. <http://doi.org/10.1111/j.1744-6570.2005.00301.x>
- Roberts, B. W., Kuncel, N. R., Shiner, R., Caspi, A., & Goldberg, L. R. (2007). The power of personality: The comparative validity of personality traits, socioeconomic status, and

cognitive ability for predicting important life outcomes. *Perspectives on Psychological Science*, 2, 313–345. <http://doi.org/10.1111/j.1745-6916.2007.00047.x>

Roberts, B. W., Lejuez, C., Krueger, R. F., Richards, J. M., & Hill, P. L. (2014). What is Conscientiousness and how can it be assessed? *Developmental Psychology*, 50, 1315–1330. <http://doi.org/10.1037/a0031109>

Ryff, C. D., & Keyes, C. L. M. (1995). The structure of psychological well-being revisited. *Journal of Personality and Social Psychology*, 69, 719–727. <http://doi.org/10.1037/0022-3514.69.4.719>

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, 1062–1075. <http://doi.org/10.1177/0146167204264292>

Schmitt, D. P., & Shackelford, T. K. (2008). Big Five traits related to short-term mating: From personality to promiscuity across 46 nations. *Evolutionary Psychology*, 6, 246–282.

Spearman, C. (1927). *The abilities of man*. Oxford, England: Macmillan.

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

Sutin, A. R., Costa, P. T. J., Chan, W., Milaneschi, Y., Eaton, W. W., Zonderman, A. B., ... Terracciano, A. (2013). I know not to, but I can't help it: Weight gain and changes in impulsivity-related personality traits. *Psychological Science*, 24, 1323–1328. <http://doi.org/10.1177/0956797612469212>

Sutin, A. R., Ferrucci, L., Zonderman, A. B., & Terracciano, A. (2011). Personality and obesity across the adult life span. *Journal of Personality and Social Psychology*, 101, 579–592. <http://doi.org/10.1037/a0024286>

Tellegen, A., & Waller, N. G. (2008). Exploring personality through test construction: Development of the Multidimensional Personality Questionnaire. In G. J. Boyle, G. Matthews, D. H. Saklofske, G. J. (Ed) Boyle, G. (Ed) Matthews, & D. H. (Ed) Saklofske (Eds.), *The SAGE handbook of personality theory and assessment, Vol 2: Personality measurement and testing*.

(pp. 261–292). Thousand Oaks, CA, US: Sage Publications, Inc.

- Terracciano, A., McCrae, R. R., Brant, L. J., & Costa, P. T., Jr. (2005). Hierarchical linear modeling analyses of the NEO-PI-R scales in the Baltimore Longitudinal Study of Aging. *Psychology and Aging, 20*, 493–506. <http://doi.org/10.1037/0882-7974.20.3.493>
- Terracciano, A., Sutin, A. R., McCrae, R. R., Deiana, B., Ferrucci, L., Schlessinger, D., ... Costa, P. T. (2009). Facets of personality linked to underweight and overweight. *Psychosomatic Medicine, 71*, 682–689. <http://doi.org/10.1097/PSY.0b013e3181a2925b>
- Tett, R. P., Steele, J. R., & Beauregard, R. S. (2003). Broad and narrow measures on both sides of the personality–job performance relationship. *Journal of Organizational Behavior, 24*, 335–356. <http://doi.org/10.1002/job.191>
- Vainik, U., Mõttus, R., Allik, J., Esko, T., & Realo, A. (2015). Are trait–outcome associations caused by scales or particular items? Example analysis of personality facets and BMI. *European Journal of Personality*. <http://doi.org/10.1002/per.2009>
- Westfall, J., Judd, C. M., & Kenny, D. A. (2015). Replicating studies in which samples of participants respond to samples of stimuli. *Perspectives on Psychological Science, 10*, 390–399. <http://doi.org/10.1177/1745691614564879>
- Weston, S. J., & Jackson, J. J. (2015). Identification of the healthy neurotic: Personality traits predict smoking after disease onset. *Journal of Research in Personality, 54*, 61–69. <http://doi.org/10.1016/j.jrp.2014.04.008>
- Wood, D., Gardner, M. H., & Harms, P. D. (2015). How functionalist and process approaches to behavior can explain trait covariation. *Psychological Review, 122*, 84–111.