

PERSPECTIVE/OPINION

Why Self-Report Measures of Self-Control and Inhibition Tasks Do Not Substantially Correlate

Lasse Wennerhold and Malte Friese

Trait self-control is often defined as the ability to inhibit dominant responses including thoughts, emotions, and behavioral impulses. Despite the pivotal role of inhibition for trait self-control, a growing body of evidence found small-to-zero correlations between self-report measures of trait self-control and behavioral inhibition tasks. These observations seem puzzling considering that both types of measures are often seen as operationalizations of the same or at least closely related theoretical constructs. Previous explanations for this non-correspondence focused on psychometric properties of the measures. Here, we discuss three further factors that may explain the empirical non-correspondence between trait self-control scales and behavioral inhibition tasks: (1) the distinction between typical and maximum performance, (2) the measurement of single versus repeated performance, and (3) differences between impulses in different domains. Specifically, we argue that a) self-report measures of trait self-control are designed to assess typical performance, and relative to these, behavioral inhibition tasks are designed to assess maximum performance; b) self-report measures of trait self-control capture central tendencies of aggregates of many different instances of behavior, whereas behavioral inhibition tasks are momentary, one-time state measures; and c) most self-report measures of trait self-control are designed to measure general, cross-domain inhibition, whereas behavioral inhibition tasks also measure narrower, domain-specific inhibition to a substantial degree. In conclusion, we argue that it is implausible to hypothesize more than a low correlation between self-report measures of trait self-control and behavioral inhibition tasks as they genuinely focus on different aspects of the theoretical construct of self-control. We also discuss the broader implications of these issues for self-control as a theoretical construct and its appropriate measurement.

Keywords: self-control; self-report measures; Self-Control Scale; inhibition tasks; executive function; typical versus maximum performance; impulses

Consider the following scenario: A researcher employs two different measurement instruments designed to measure the same construct. To her surprise, she ends up with a low correlation between the two measures. A different researcher does the same thing with slightly different measures – again, almost no correlation between the measures. How is this possible? These (non-)correlations require an explanation.

In recent years, a situation akin to the one described in the previous paragraph has emerged in the field of self-control. A growing body of evidence found small-to-zero correlations between self-report measures of trait self-control and behavioral inhibition tasks (Allom, Panetta, Mullan, & Hagger, 2016; Duckworth & Kern, 2011; Eisenberg et al., 2019; Nęcka, Gruszka, Orzechowski, Nowak, & Wójcik, 2018; Saunders, Milyavskaya, Etz,

Randles, & Inzlicht, 2018). For example, Saunders et al. (2018) published a series of studies suggesting that trait self-control and inhibitory control are not meaningfully associated. Trait self-control was assessed with the most prominent self-report measure in the field (i.e., the Self-Control Scale, Tangney, Baumeister, & Boone, 2004); inhibitory control was assessed with two established executive function tasks (i.e., the Stroop task, Stroop, 1935, and the Flanker task, Eriksen & Eriksen, 1974) that are commonly used to assess inhibitory control (e.g., Diamond, 2013). An internal Bayesian meta-analysis across 5 studies (overall $N > 2,600$) revealed a null relationship between the two types of measures. Similar results emerged for other constructs closely related to trait self-control, for example, conscientiousness and impulsivity (Edmonds, Bogg, & Roberts, 2009; Eisenberg et al., 2019; Enkavi et al., 2019; Fleming, Heintzelmann, & Bartholow, 2016; Frey, Pedroni, Mata, Rieskamp, & Hertwig, 2017; Stahl et al., 2014).

These empirical observations seem puzzling. After all, self-report measures of trait self-control

and behavioral inhibition tasks are often seen as operationalizations of the same or at least closely related theoretical constructs. They are designed with the intention to capture the ability of a person to inhibit dominant responses including thoughts, emotions, and behavioral impulses (Tangney et al., 2004). This ability is classically considered the central definitional aspect of the construct of self-control (e.g., Baumeister, 2014; Duckworth & Kern, 2011; Hofmann, Schmeichel, & Baddeley, 2012). Thus, it would seem appropriate to expect moderate or even strong correlations between these measures. When we talk about self-control in this article, we also refer to this conceptualization of the construct.

Saunders et al. (2018) discuss three factors that might contribute to this seeming conundrum. First, the Self-Control Scale (Tangney et al., 2004) might capture other things than inhibition alone (e.g., items like “I am able to work effectively toward long-term goals”, “I engage in healthy practices”) while the focus of behavioral inhibition tasks is much narrower. Second, behavioral inhibition tasks are often selected for little between-subject variability, which limits their ability to assess between-person differences and correlate with other individual difference measures (reliability paradox, Enkavi et al., 2019; Hedge, Powell, & Sumner, 2018; Rouder, Kumar, & Haaf, 2019). Third, previous meta-analytic estimates of the relationship between self-report measures of trait self-control and behavioral inhibition tasks (e.g., Duckworth & Kern, 2011) may have overestimated the relationship due to publication bias (which does not explain the low correlations, but the surprise about them).

We agree with all three arguments Saunders et al. (2018) consider to explain the observed negligible correlations between measures of trait self-control and inhibition. In the present article, we offer three further factors not addressed by Saunders et al. (2018) that may explain this empirical non-correspondence. From our perspective, considering these factors is important for two reasons: First, they help to explain a series of seemingly surprising empirical findings (Allom et al., 2016; Duckworth & Kern, 2011; Eisenberg et al., 2019; Nečka et al., 2018; Saunders et al., 2018). Second, despite them being rarely considered in the self-control literature, we believe that these issues are of broader relevance as they help to distinguish between trait self-control as a construct and its operationalization by measurement instruments. The first issue is the distinction between typical and maximum performance; the second refers to the measurement of single versus repeated performance; the third relates to differences between impulses in different domains.

Because of these issues, we argue that it is not plausible to hypothesize more than a low correlation between self-report measures of trait self-control and behavioral inhibition tasks, *even if* a) the respective self-report measure of trait self-control would measure inhibitory processes alone and b) the behavioral inhibition tasks would show high retest reliability in measuring between-person differences. Putting psychometric issues aside, both measurement approaches genuinely focus on different aspects of the theoretical construct of

self-control. While most self-report measures of trait self-control are designed to assess trait-like typical inhibitory performance that is repeatedly shown across a broad range of impulses from different domains, behavioral inhibition tasks are designed to measure ability-like maximum inhibitory performance shown on single occasions for specific kinds of impulses.

Note that we use the wording “self-report measures of trait self-control” and “behavioral inhibition tasks” to refer to methodological categories of operationalizations: self-report questionnaires on the one hand and performance tasks on the other hand. We do this to stress the point that, although we refer to the Self-Control Scale and the Stroop/Flanker tasks as prominent examples of their respective methodological categories following the example of Saunders et al. (2018), most issues raised in this article do not only pertain to the Self-Control Scale and the Stroop/Flanker tasks specifically, but their respective methodological categories more generally. When discussing issues that refer to the Self-Control Scale or the Stroop/Flanker tasks as specific operationalizations, we name them directly as “the Self-Control Scale” or “the Stroop/Flanker task”. In general, we refer to the level of self-control measurement when not directly stating that we refer to the theoretical construct level of self-control. We do not discuss which methodological approach towards the measurement of self-control, self-report questionnaires or performance tasks, may generally be preferable as a benchmark measure of self-control.

Typical versus maximum self-control performance

In industrial and organizational psychology, there exists the well-established distinction between typical and maximum performance (Sackett, Zedeck, & Fogli, 1988). Typical performance refers to the *tendency* to perform relatively consistently at some level across different situations over a prolonged period of time. By contrast, maximum performance refers to the *ability* to perform at the highest possible level on a specific occasion. Although typical and maximum performance are not independent, they correlate only moderately (for a meta-analysis, see Beus & Whitman, 2012).

The distinction between typical and maximum performance nicely maps onto the distinction between self-report measures of trait self-control and behavioral inhibition tasks. The instruction of the Self-Control Scale (Tangney et al., 2004) asks respondents to “[...] indicate how much each of the following statements reflects how you *typically* are” (emphasis added, p. 323), thus clearly asking about respondents’ typical behavior rather than their maximum ability. This also pertains to the wording of many of the scale’s items that make clear that what respondents are asked for is typical behavior rather than maximum ability (e.g., “I am lazy”). Similar observations apply to other self-report measures of trait self-control; they are designed to assess typical thoughts, feelings, and behaviors of respondents.

Consistent with the notion that the Self-Control Scale assesses typical rather than maximum performance,

growing evidence suggests that the plentiful associations of the Self-Control Scale with desirable life outcomes are to a large extent due to beneficial stable habits, not due to effortful inhibition in particular situations (e.g., de Ridder et al., 2012; Galla & Duckworth, 2015; Grund & Carstens, 2019; Hofmann, Baumeister, Förster, & Vohs, 2012). In other words, it is not mainly individual differences in maximum performance, but in typical performance that differentiates between respondents on this scale.

If we assume a continuum between typical performance on one end and maximum performance on the other end, behavioral inhibition tasks are clearly designed to assess behavior more on the maximum performance end of the continuum compared to self-report instruments such as the Self-Control Scale. In behavioral inhibition tasks, respondents are instructed to avoid making errors and/or to try being fast in the task that is to follow. Nothing suggests to participants that what the researchers seek to measure is something akin to how they typically behave. Rather, the context makes it clear that what counts is a performance that is as good as possible.

Admittedly, many people might not be maximally motivated and thus might not show their *absolute* maximum performance on behavioral inhibition tasks when they take part in a scientific study. Thus, provided strong enough incentives it may be possible to improve performance beyond baseline levels. The point that we make here is that there is a pronounced difference between self-report measures of trait self-control and behavioral inhibition tasks. Self-report instruments are designed to measure something clearly more on the side of typical performance and behavioral inhibition tasks are designed to measure something clearly more on the side of maximum performance.

The distinction between typical and maximum performance has not been picked up in the self-control literature to a great extent. One exception is a study by Freudenthaler and Neubauer (2007) in the domain of emotion management, an important aspect of self-control (see also Nečka et al., 2018). Typical and maximum emotion management performance were measured between participants. Their correspondence could therefore not be assessed. However, the authors found that self-reported typical emotion management performance was less optimal than maximum performance. In addition, self-reported typical emotion management varied more strongly between persons than maximum performance. In other words, most participants were aware of adequate emotion management strategies that should optimally be used (maximum performance), but not all of them reported typically employing them in the relevant situations. This suggests that typical performance might be a more reliable indicator of between-person differences compared to maximum performance. This corresponds to the discussion about the reliability paradox mentioned previously, that is, the selection of behavioral inhibition tasks and other executive function tasks for little instead of ample between-person differences (Enkavi et al., 2019; Hedge et al., 2018).

Taken together, the distinction between typical and maximum performance helps to explain low correlations between self-report measures of trait self-control and behavioral inhibition tasks because (1) self-report measures such as the Self-Control Scale (Tangney et al., 2004) are designed to assess typical performance, and behavioral inhibition tasks are designed to assess maximum performance, (2) typical and maximum performance are conceptually different and empirically only modestly related, and (3) maximum performance measures might show comparatively little between-person variability, limiting their ability to correlate with other individual differences. Note that we are not claiming that the distinction between typical and maximum performance is *the* factor that explains low correlations between self-report measures of trait self-control and behavioral inhibition tasks. In fact, the moderate associations between the two types of performance (Beus & Whitman, 2012) would still predict a discernable correlation. Nevertheless, this distinction is one factor worth considering to understand why self-report measures of trait self-control and behavioral inhibition tasks show small-to-zero correlations.

We believe that the implications of the distinction between typical and maximum performance go beyond the explanation of low correlations between self-report measures of trait self-control and behavioral inhibition tasks. One implication refers to the measurement of trait self-control. A seminal definition describes traits as “[...] individual differences in tendencies to show consistent patterns of thoughts, feelings, and actions” (McCrae & Costa, 2003, p. 25, see also Roberts, 2009; Roberts & Jackson, 2008). The typical versus maximum distinction highlights that having the ability to perform at a certain maximum level does not imply the tendency to consistently live up to this ability. Thus, when measurement instruments seek to assess *trait* self-control they should focus on relatively consistent tendencies to show self-control, not on maximum ability. In other cases, researchers may specifically conceptualize self-control as an ability or specify different dimensions of self-control, some of which might refer to ability aspects of the construct and others to trait-like aspects, and measure them separately.

A second implication refers to associations between measures of self-control and outcome variables. Behaving in a self-controlled way is not particularly hard for most situations prototypically seen as self-control dilemmas. Yes, some impulses are stronger than others, but in principle, most persons are *able* to eat an apple instead of a chocolate bar or work instead of checking social media in any specific situation. Whether or not they do so will likely have little impact on their overall success in life. That is not to say that single acts of behavior may not have profound impacts on people’s lives. However, desirable life outcomes that are associated with trait self-control such as academic achievement, financial wealth, health, or stable social relationships (Moffitt et al., 2011; Tangney et al., 2004), are usually the result of “doing the right thing” most of the time over extended periods of time, not

in single situations. In other words, what differentiates people with high versus low trait self-control and helps them to achieve a multitude of desirable life outcomes over months, years, and decades is that people with high trait self-control tend to *typically* act in a more self-controlled manner.

Single versus repeated performance

The observation that relative to self-report instruments such as the Self-Control Scale behavioral inhibition tasks assess something more akin to maximum performance can partly explain the low correspondence between these variables. However, even if one would assume that, on average, participants in scientific studies show so little motivation to perform well on behavioral inhibition tasks that these tasks do not work as measures of maximum performance at all, they would likely make relatively poor trait measures. The reason is that any measurement of inhibition would remain a momentary, one-time, state measure of inhibitory performance that would be heavily influenced by situational factors and likely be unreliable as an indicator of this person's typical performance. By contrast, self-report instruments that ask about typical behavior like the Self-Control scale provoke responses that are central tendencies of aggregates of many different instances of behavior and are therefore more likely to reliably grasp trait levels.

In his seminal contributions to the person-situation debate, Fleeson (2001, 2004) discovered that personality traits are characterized by strong intraindividual variability. That is, a moderately self-controlled person does not behave in a moderately self-controlled way all the time (**Figure 1A**). Instead, there are many instances in which this person behaves in a more or less self-controlled way than at a moderate level (**Figure 1B**).¹ What justifies calling this person moderately self-controlled is that *on average* across many different situations the central

tendency of this person's behavior is relatively stable at a moderate level of self-control (**Figure 1C**), compared to that of others. For example, the average self-control level of a person during a certain time period (e.g., one week) tends to be similar to this person's average self-control level during another time period (e.g., the following week). This led Fleeson to conclude that traits are density distributions of states such as those depicted in **Figure 1**. Thus, there is strong variability within persons, but also high stability of the central tendencies between persons. This intriguing insight effectively resolved a good part of the person-situation debate that kept social and personality psychologists busy for decades (Fleeson & Jayawickreme, 2015).

We regard inhibition as a personality trait. **Figure 1B** illustrates that a one-shot behavioral measurement of inhibition is unlikely to reflect the central tendency of a person. For behavioral inhibition tasks to indicate typical performance they would need to be applied multiple times in different circumstances, ideally in participants' daily lives via ecological momentary assessment (see Shiffman, Stone, & Hufford, 2008, for an overview) to increase external validity. The result would be a density distribution of inhibition states, the central tendency of which would be a more reliable indicator of typical inhibition (that is, the trait). This indicator should correlate more substantially with any other measure indicating typical inhibition performance than single-shot measures of the same construct.

Inhibition of different kinds of impulses

Another factor that may contribute to explaining the null relationship between self-report measures of trait self-control and behavioral inhibition tasks is that both measures differ in the range of impulses that are inhibited in different domains they focus on. Specifically, we argue that self-report measures of trait self-control typically



Figure 1: Each graph depicts the number of times a hypothetical person acted at each level of self-control. The graph on the left (A) depicts the density distribution of self-controlled behavior with on average moderate level of self-control and relatively low intraindividual variability (i.e., this fictitious person almost always behaved in a moderately self-controlled way). The graph in the middle (B) also depicts the density distribution of a self-controlled behavior with on average moderate level of self-control, this time with relatively high intraindividual variability (i.e., this fictitious person often behaves considerably more or less self-controlled than on a moderate level). Fleeson (2001) found that actual distributions more resemble **Figure 1B** than **Figure 1A**. In **Figure 1C**, each point in this graph represents one person's average level of self-control in two different time periods (e.g., one week). The work by Fleeson (2001, 2004) suggests that how self-controlled a person acts on average in one time period is highly similar to how self-controlled the person acts on average in another time period. Figure adapted from Fleeson (2004).

measure general, cross-domain inhibition, whereas behavioral inhibition tasks measure also narrower, domain-specific inhibition to a substantial degree.

The role of potential qualitative differences of impulses in different domains is rarely discussed in the self-control literature (e.g., an impulse to eat unhealthy food versus an impulse to insult someone in an argument). As a consequence, a naive reader of the literature might assume that there are few qualitative differences between impulses across domains and that people are typically able to inhibit impulses in different domains to roughly the same extent. Domain-general measures of trait self-control – like the Self-Control Scale (Tangney et al., 2004) – predict a wide range of desirable life outcomes across different domains (e.g., de Ridder et al., 2012; Tangney et al., 2004). It is therefore indeed plausible that there are causes of self-control performance that might be common across domains. At the same time, a growing literature advances theoretical arguments and/or provides empirical evidence suggesting that there are noteworthy differences between domain-general and domain-specific self-control (de Ridder et al., 2012; Duckworth & Tsukayama, 2015; Eisenberg et al., 2019; Haws, Davis, & Dholakia, 2016; Roberts, Lejuez, Krueger, Richards, & Hill, 2014). For example, Haws et al. (2016) report only moderate correlations between the Self-Control Scale (Tangney et al., 2004) and domain-specific adaptations of this scale for spending and eating behavior. To be successful in a given domain, people likely need domain-general self-control plus domain-specific skills.

In their meta-analysis of the associations of different self-report measures of trait self-control with behavior in various life domains, de Ridder et al. (2012) found substantial variability of these relationships across domains. Thus, it seems plausible that people differ in their domain-specific self-control across domains. This could be due to a) differences between impulses in different domains that make – on average – impulses feel stronger in one domain than in another (i.e., a bottom-up process) or b) due to differences in the capacity to inhibit impulses across different domains (i.e., a top-down process). The Self-Control Scale is conceptualized as an instrument to indicate the strength of the inhibitory top-down process and does not distinguish between these different aspects.

Admittedly, it is a matter of debate to what extent the Self-Control Scale measures the inhibition of impulses alone, whether domain-general or domain-specific. Likely, it measures also other, and even a broad array of processes. This is one of the points Saunders et al. (2018) made to explain the zero-correlation between the scale and behavioral inhibition tasks. It is also currently unknown to what extent the inhibition of impulses per se is conducive to the relationships of the Self-Control Scale with real-life outcomes. The point here is that *even if* the Self-Control Scale measured general, cross-domain inhibition alone (which is what it is intended to do) we should not expect particularly strong correlations with behavioral inhibition tasks (e.g., Stroop task, Flanker task).

We noted doubts about the validity of the Self-Control scale as a (more or less) pure measure of domain-general

inhibition. Similarly, the validity of behavioral inhibition tasks as pure measures of inhibition can be questioned. Whether performance on these tasks is determined by inhibitory control alone or other processes as well, for example, selective attention, is not entirely clear (Cohen, Dunbar, & McClelland, 1990; Egnor & Hirsch, 2005). This notwithstanding, these tasks arguably measure inhibition to a substantial degree and are commonly used as measures of inhibition. Thus, one should expect at least a moderate correlation of these measures with a self-report scale that measures inhibition if both approaches captured similar inhibitory processes.

Behavioral inhibition tasks are meant to assess domain-general inhibition as well. It is not clear, however, to what extent they actually achieve this goal. We argue that these tasks substantially measure the inhibition of rather specific kinds of impulses as well. For instance, in the Stroop task reading a written word evokes an impulse to indicate the lexical meaning of that word. As respondents are given a different task, namely to indicate the physical color of a written word, they have to overcome the impulse to refer to the lexical meaning when physical color and lexical meaning are incompatible. This task clearly involves some kind of inhibition. However, we deem it implausible to assume that inhibiting to name the lexical meaning of a word in split-second decision making is qualitatively identical to the kind of inhibition that is required when inhibiting other impulses from other domains (e.g., resisting a palatable chocolate bar or resisting to insult another person in a heated discussion). In other words, although the Stroop task (and other behavioral inhibition tasks for that matter) may appear to be domain-general measures of inhibition due to their affectively cold and abstract content and design, it likely captures substantial aspects of inhibitory control that are specific to these particular contents and that may not easily generalize to the inhibition of impulses from real life domains. The more task-specific inhibition is captured by these tasks, the more their capacity to predict outcomes across domains would be compromised. Indeed, this reasoning is in line with recent empirical evidence finding very low correlations between executive function measures and various life outcomes (e.g., Eisenberg et al., 2019).

One possible solution to this problem is to assess a variety of different inhibition tasks and model a latent variable that indicates general, cross-domain inhibition (e.g., Miyake & Friedman, 2012; Miyake et al., 2000). The resulting latent variable would represent the shared variance between the different measures and thereby reflect more general inhibitory ability. In line with our reasoning, studies that model a latent inhibition variable based on different inhibition tasks (e.g., Stroop, Stop-Signal, Antisaccade) often find relatively low standardized factor loadings ranging in the .30s to .50s, indicating that despite their similarities these tasks assess substantial task-specific aspects (Miyake & Friedman, 2012; Miyake et al., 2000).

To the extent that the Self-Control Scale or other self-report measures of trait self-control capture general inhibitory ability, a latent variable approach would

increase measurement correspondence between these scales and (the latent variable of) inhibition. Measurement correspondence has been widely discussed in other fields. For example, the seminal correspondence principle by Ajzen and Fishbein (1977) suggests that the relation between measures of attitudes and behavior increases with increasing correspondence between the measurement instruments across different entities (e.g., target, action, context, time). Strong attitude-behavior relations can only be expected when there is high correspondence between at least some of these entities. Applied to the present context, a higher correspondence between the types of inhibition – general versus domain-specific – assessed by self-report measures and the indicator of inhibition derived from behavioral inhibition tasks should increase their empirical correlation.

Taken together, we argue that domain-general self-report measures of trait self-control (e.g., the Self-Control Scale), and behavioral inhibition tasks (e.g., Stroop, Flanker) might measure the inhibition of different types of impulses varying in specificity. Self-report measures typically measure the inhibition of a broad array of impulses while behavioral inhibition tasks might measure (among other things) domain-general inhibition, but also to a substantial degree the inhibition of narrower task-specific impulses that are not representative of impulses encountered in various domains of daily life. Again, we do not claim that this issue is *the* factor explaining the lack of correspondence between both types of measures. It is plausible that there are common aspects of inhibition across domains, and this should lead to a moderate correlation if this factor would be the only relevant issue. However, the issue might be one of several factors explaining the empirical non-correlation between self-report measures of trait self-control and behavioral inhibition tasks.

Conclusion

We discussed three issues that might explain the empirical small-to-zero relationship between self-report measures of trait self-control and behavioral inhibition tasks found by a growing number of studies (e.g., Saunders et al., 2018; Nęcka et al., 2018): the distinction between typical versus maximum performance, the distinction between single versus repeated performance, and the relevance of considering different kinds of impulses. Bearing these issues in mind, we argue that it is implausible to hypothesize more than a low correlation between self-report measures of trait self-control and behavioral inhibition tasks *even if* a) the respective self-report measure of trait self-control would measure inhibitory processes alone and b) the inhibition-related measures of executive function would show high retest reliability in measuring between-person differences. Beyond psychometric issues, both approaches genuinely focus on distinct facets of the theoretical construct of self-control. Most self-report measures of trait self-control are designed to assess trait-like typical inhibitory performance that is repeatedly shown across a broad range of impulses from different domains. Behavioral inhibition tasks,

by contrast, are typically designed to measure ability-like maximum inhibitory performance shown on single occasions for more specific kinds of impulses. Future theoretical and empirical research should examine more closely which of the discussed (and possibly additional) factors contribute to which extent to the degree of empirical (non-)correspondence of different measures of self-control, helping the field to understand both the theoretical nature and suitable measurement approaches of the construct better.

Note

- ¹ Fleeson (2001) did not examine this idea with trait self-control specifically, but with other traits like agreeableness, extraversion, or conscientiousness (the latter being conceptually closely related to trait self-control).

Acknowledgements

The authors thank Blair Saunders, an anonymous reviewer, and the handling editor Simine Vazire for stimulating comments and fruitful advice on how to improve and expand this manuscript.

Competing Interests

The authors have no competing interests to declare.

Author Contributions

LW and MF conceived the manuscript. LW drafted the initial manuscript. LW and MF revised the manuscript.

References

- Ajzen, I., & Fishbein, M. (1977). Attitude-behavior relations: A theoretical analysis and review of empirical research. *Psychological Bulletin*, *84*, 888–918. DOI: <https://doi.org/10.1037/0033-2909.84.5.888>
- Allom, V., Panetta, G., Mullan, B., & Hagger, M. S. (2016). Self-report and behavioural approaches to the measurement of self-control: Are we assessing the same construct? *Personality and Individual Differences*, *90*, 137–142. DOI: <https://doi.org/10.1016/j.paid.2015.10.051>
- Baumeister, R. F. (2014). Self-regulation, ego depletion, and inhibition. *Neuropsychologia*, *65*, 313–319. DOI: <https://doi.org/10.1016/j.neuropsychologia.2014.08.012>
- Beus, J. M., & Whitman, D. S. (2012). The relationship between typical and maximum performance: A meta-analytic examination. *Human Performance*, *25*, 355–376. DOI: <https://doi.org/10.1080/08959285.2012.721831>
- Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: A parallel distributed processing account of the Stroop effect. *Psychological Review*, *97*, 332–361. DOI: <https://doi.org/10.1037/0033-295X.97.3.332>
- de Ridder, D. T. D., Lensvelt-Mulders, G., Finkenauer, C., Stok, F. M., & Baumeister, R. F. (2012). Taking stock of self-control: A meta-analysis of how trait self-control relates to a wide range of behaviors.

- Personality and Social Psychology Review*, 16, 76–99. DOI: <https://doi.org/10.1177/1088868311418749>
- Diamond, A.** (2013). Executive functions. *Annual Review of Psychology*, 64, 135–168. DOI: <https://doi.org/10.1146/annurev-psych-113011-143750>
- Duckworth, A. L., & Kern, M. L.** (2011). A meta-analysis of the convergent validity of self-control measures. *Journal of Research in Personality*, 45, 259–268. DOI: <https://doi.org/10.1016/j.jrp.2011.02.004>
- Duckworth, A. L., & Tsukayama, E.** (2015). Domain specificity in self-control. In C. B. Miller, R. M. Furr, A. Knobel, & W. Fleeson (Eds.), *Character: New directions from philosophy, psychology, and theology* (pp. 393–411). New York: Oxford University Press. DOI: <https://doi.org/10.1093/acprof:oso/9780190204600.003.0019>
- Edmonds, G. W., Bogg, T., & Roberts, B. W.** (2009). Are personality and behavioral measures of impulse control convergent or distinct predictors of health behaviors? *Journal of Research in Personality*, 43, 806–814. DOI: <https://doi.org/10.1016/j.jrp.2009.06.006>
- Egner, T., & Hirsch, J.** (2005). Cognitive control mechanisms resolve conflict through cortical amplification of task-relevant information. *Nature Neuroscience*, 8, 1784–1790. DOI: <https://doi.org/10.1038/nn1594>
- Eisenberg, I. W., Bissett, P. G., Enkavi, A. Z., Li, J., MacKinnon, D. P., Marsch, L. A., & Poldrack, R. A.** (2019). Uncovering the structure of self-regulation through data-driven ontology discovery. *Nature Communications*, 10, 2319. DOI: <https://doi.org/10.1038/s41467-019-10301-1>
- Enkavi, A. Z., Eisenberg, I. W., Bissett, P. G., Mazza, G. L., Mackinnon, D. P., & Marsch, L. A.** (2019). Large-scale analysis of test – retest reliabilities of self-regulation measures. *Proceedings of the National Academy of Sciences of the United States of America*, 116, 5472–5477. DOI: <https://doi.org/10.1073/pnas.1818430116>
- Eriksen, B. A., & Eriksen, C. W.** (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, 16, 143–149. DOI: <https://doi.org/10.3758/BF03203267>
- Fleeson, W.** (2001). Toward a structure- and process-integrated view of personality: traits as density distributions of states. *Journal of Personality and Social Psychology*, 80, 348–360. DOI: <https://doi.org/10.1111/jopy.12163>
- Fleeson, W.** (2004). Moving personality beyond the person-situation debate: The challenge and the opportunity of within-person variability. *Current Directions in Psychological Science*, 13, 83–87. DOI: <https://doi.org/10.1111/j.0963-7214.2004.00280.x>
- Fleeson, W., & Jayawickreme, E.** (2015). Whole trait theory. *Journal of Research in Personality*, 56, 82–92. DOI: <https://doi.org/10.1016/j.jrp.2014.10.009>
- Fleming, K. A., Heintzelman, S. J., & Bartholow, B. D.** (2016). Specifying associations between conscientiousness and executive functioning: Mental set shifting, not prepotent response inhibition or working memory updating. *Journal of Personality*, 84, 1011–1027. DOI: <https://doi.org/10.1037//0022-3514.80.6.1011>
- Freudenthaler, H. H., & Neubauer, A. C.** (2007). Measuring emotional management abilities: Further evidence of the importance to distinguish between typical and maximum performance. *Personality and Individual Differences*, 42, 1561–1572. DOI: <https://doi.org/10.1016/j.paid.2006.10.031>
- Frey, R., Pedroni, A., Mata, R., Rieskamp, J., & Hertwig, R.** (2017). Risk preference shares the psychometric structure of major psychological traits. *Science Advances*, 3, e1701381. DOI: <https://doi.org/10.1126/sciadv.1701381>
- Galla, B. M., & Duckworth, A. L.** (2015). More than resisting temptation: Beneficial habits mediate the relationship between self-control and positive life outcomes. *Journal of Personality and Social Psychology*, 109, 508–525. DOI: <https://doi.org/10.1037/pspp0000026>
- Grund, A., & Carstens, C. A.** (2019). Self-control motivationally reconsidered: “Acting” self-controlled is different to “being good” at self-control. *Motivation and Emotion*, 43, 63–81. DOI: <https://doi.org/10.1007/s11031-018-9721-3>
- Haws, K. L., Davis, S. W., & Dholakia, U. M.** (2016). Control over what? Individual differences in general versus eating and spending self-control. *Journal of Public Policy & Marketing*, 35, 37–57. DOI: <https://doi.org/10.1509/jppm.14.149>
- Hedge, C., Powell, G., & Sumner, P.** (2018). The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. *Behavior Research Methods*, 50, 1166–1186. DOI: <https://doi.org/10.3758/s13428-017-0935-1>
- Hofmann, W., Baumeister, R. F., Förster, G., & Vohs, K. D.** (2012). Everyday temptations: An experience sampling study of desire, conflict, and self-control. *Journal of Personality and Social Psychology*, 102, 1318–1335. DOI: <https://doi.org/10.1037/a0026545>
- Hofmann, W., Schmeichel, B. J., & Baddeley, A. D.** (2012). Executive functions and self-regulation. *Trends in Cognitive Sciences*, 16, 174–180. DOI: <https://doi.org/10.1016/j.tics.2012.01.006>
- McCrae, R. R., & Costa, P. T.** (2003). *Personality in adulthood: A five-factor theory perspective* (2nd Ed.). New York, NY, US: Guilford Press. DOI: <https://doi.org/10.4324/9780203428412>
- Miyake, A., & Friedman, N. P.** (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science*, 21, 8–14. DOI: <https://doi.org/10.1177/0963721411429458>
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D.** (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent

- variable analysis. *Cognitive Psychology*, 41, 49–100. DOI: <https://doi.org/10.1006/cogp.1999.0734>
- Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H., ... & Sears, M. R.** (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proceedings of the National Academy of Sciences*, 108, 2693–2698. DOI: <https://doi.org/10.1073/pnas.1010076108>
- Necka, E., Gruszka, A., Orzechowski, J., Nowak, M., & Wójcik, N.** (2018). The (in)significance of executive functions for the trait of self-control: A psychometric study. *Frontiers in Psychology*, 9, 1139. DOI: <https://doi.org/10.3389/fpsyg.2018.01139>
- Roberts, B. W.** (2009). Back to the future: Personality and assessment and personality development. *Journal of Research in Personality*, 43, 137–145. DOI: <https://doi.org/10.1016/j.jrp.2008.12.015>
- Roberts, B. W., & Jackson, J. J.** (2008). Sociogenic personality psychology. *Journal of Personality*, 76, 1523–1544. DOI: <https://doi.org/10.1111/j.1467-6494.2008.00530.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. DOI: <https://doi.org/10.1037/a0031109>
- Rouder, J., Kumar, A., & Haaf, J. M.** (2019, March 25). Why most studies of individual differences with inhibition tasks are bound to fail. Retrieved from psyarxiv.com/3cjr5. DOI: <https://doi.org/10.31234/osf.io/3cjr5>
- Sackett, P. R., Zedeck, S., & Fogli, L.** (1988). Relations between measures of typical and maximum job performance. *Journal of Applied Psychology*, 73, 482–486. DOI: <https://doi.org/10.1097/BRS.0b013e3181da3737>
- Saunders, B., Milyavskaya, M., Etz, A., Randles, D., & Inzlicht, M.** (2018). Reported self-control is not meaningfully associated with inhibition-related executive function: A Bayesian analysis. *Collabra: Psychology*, 4, 39. DOI: <https://doi.org/10.1525/collabra.134>
- Shiffman, S., Stone, A. A., & Hufford, M. R.** (2008). Ecological momentary assessment. *Annual Review of Clinical Psychology*, 4, 1–32. DOI: <https://doi.org/10.1146/annurev.clinpsy.3.022806.091415>
- Stahl, C., Voss, A., Schmitz, F., Nuszbaum, M., Tüscher, O., Lieb, K., & Klauer, K. C.** (2014). Behavioral components of impulsivity. *Journal of Experimental Psychology: General*, 143, 850–886. DOI: <https://doi.org/10.1037/a0033981>
- Stroop, R.** (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18, 643–661. DOI: <https://doi.org/10.1037/h0054651>
- Tangney, J. P., Baumeister, R. F., & Boone, A. L.** (2004). High self-control predicts good adjustment, less pathology, better grades, and interpersonal success. *Journal of Personality*, 72, 271–324. DOI: <https://doi.org/10.1111/j.0022-3506.2004.00263.x>

Peer review comments

The author(s) of this paper chose the Open Review option, and the peer review comments can be downloaded at: <http://doi.org/10.1525/collabra.276.pr>

How to cite this article: Wennerhold, L., & Friese, M. (2020). Why Self-Report Measures of Self-Control and Inhibition Tasks Do Not Substantially Correlate. *Collabra: Psychology*, 6(1): 9. DOI: <https://doi.org/10.1525/collabra.276>

Senior Editor: Simine Vazire

Editor: Simine Vazire

Submitted: 25 July 2019 **Accepted:** 03 January 2020 **Published:** 28 January 2020

Copyright: © 2020 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.