JOURNAL OF PERSONALITY ASSESSMENT
2021, VOL. 103, NO. 2, 238–245
https://doi.org/10.1080/00223891.2020.1732994

Correlation Database of 60 Cross-Disciplinary Surveys and Cognitive Tasks Assessing Self-Regulation

Gina L. Mazza[1], Heather L. Smyth[2], Patrick G. Bissett[3], Jessica R. Canning[4], Ian W. Eisenberg[3], A. Zeynep Enkavi[3], Oscar Gonzalez[5], Sunny Jung Kim[6], Stephen A. Metcalf[7], Felix Muniz[2], William E. Pelham III[2], Emily A. Scherer[8], Matthew J. Valente[9], Haiyi Xie[7], Russell A. Poldrack[3], Lisa A. Marsch[7,10], and David P. MacKinnon[2]

[1]Department of Health Sciences Research, Mayo Clinic, Scottsdale, AZ; [2]Department of Psychology, Arizona State University, Tempe, AZ; [3]Department of Psychology, Stanford University, Stanford, CA; [4]Department of Psychology, University of Washington, Seattle, WA; [5]Department of Psychology and Neuroscience, University of North Carolina at Chapel Hill, Chapel Hill, NC; [6]Department of Health Behavior and Policy, School of Medicine, Virginia Commonwealth University, Richmond, VA; [7]Department of Biomedical Data Science, Dartmouth College, Lebanon, NH; [8]Geisel School of Medicine, Dartmouth College, Lebanon, NH; [9]Department of Psychology, Florida International University, Miami, FL; [10]Department of Psychiatry, Dartmouth College, Lebanon, NH

ABSTRACT

Self-regulation is studied across various disciplines, including personality, social, cognitive, health, developmental, and clinical psychology; psychiatry; neuroscience; medicine; pharmacology; and economics. Widespread interest in self-regulation has led to confusion regarding both the constructs within the nomological network of self-regulation and the measures used to assess these constructs. To facilitate the integration of cross-disciplinary measures of self-regulation, we estimated product-moment and distance correlations among 60 cross-disciplinary measures of self-regulation (23 self-report surveys, 37 cognitive tasks) and measures of health and substance use based on 522 participants. The correlations showed substantial variability, though the surveys demonstrated greater convergent validity than did the cognitive tasks. Variables derived from the surveys only weakly correlated with variables derived from the cognitive tasks ($M = .049$, range = .000 to .271 for the absolute value of the product-moment correlation; $M = .085$, range = .028 to .241 for the distance correlation), thus challenging the notion that these surveys and cognitive tasks measure the same construct. We conclude by outlining several potential uses for this publicly available database of correlations.

ARTICLE HISTORY
Received 27 March 2019
Accepted 19 January 2020

Self-regulation is important because of its potential role in the development and maintenance of many behaviors. As recently defined by the National Institutes of Health (NIH), "self-regulation refers to the process of managing emotional, motivational and cognitive resources to align mental states and behavior with our goals" (NIH, 2015). Constructs within the nomological network of self-regulation have been implicated in a variety of risky health behaviors and outcomes, including poor diet; physical inactivity; alcohol, tobacco, and other substance use problems; risky sexual behaviors; risky driving; and longevity (Bickel, Odum, & Madden, 1999; Bogg & Roberts, 2004; Kern & Friedman, 2008; Moffitt et al., 2011). Modifiable behaviors such as these substantially affect health and account for approximately 40% of the risk associated with preventable premature deaths in the United States (NIH, 2015). Thus, interest lies in the characteristics of people who can and cannot successfully initiate and sustain behavior change. As part of the NIH's Science of Behavior Change initiative, our project focuses on better understanding and measuring self-regulation.

Accurately measuring self-regulation and identifying potentially modifiable facets of self-regulation are crucial to the development of more effective interventions. However, widespread interest in self-regulation has led to "conceptual confusion and measurement mayhem" (Morrison & Grammer, 2016, p. 327). As summarized by an NIH report, the nomological network of self-regulation "encompasses a wide range of behavioral and psychological constructs and processes, including, but not limited to: conscientiousness, self-control, response inhibition, impulsivity/impulse control, behavioral disinhibition, temporal discounting, emotion regulation, cognitive control (including goal selection, updating, representation and maintenance; response selection, inhibition or suppression; and performance or conflict monitoring), cognitive/emotional homeostasis, effort modulation, and flexible adaptation" (NIH, 2015; also see Nigg, 2017, Table 1). Constructs within the nomological network of self-regulation are studied across various disciplines, including personality, social, cognitive, health, developmental, and clinical psychology; psychiatry; neuroscience; medicine; pharmacology; and economics (Nigg, 2017). However,

CONTACT Gina L. Mazza Mazza.Gina@mayo.edu Department of Health Sciences Research, Mayo Clinic, Scottsdale, AZ 85259.

limited cross-talk among researchers from these disciplines has led to discipline-specific terms and measures for these constructs.

Given the "extraordinary diversity" (Duckworth & Kern, 2011, p. 259) in how self-regulation is operationalized and evaluated across disciplines, several calls for clarification and cross-disciplinary integration have been made (e.g., Duckworth & Kern, 2011; Zhou, Chen, & Main, 2012; Welsh & Peterson, 2014; Morrison & Grammer, 2016; Nigg, 2017; Eisenberg, 2017). These calls have applied to both the *constructs* within the nomological network of self-regulation and the *measures* used to assess these constructs. In this paper, we focus on the *measures* while remaining agnostic to competing theories within and across disciplines regarding how to operationalize self-regulation. Whereas others have provided conceptual clarification based on theory and expertise (e.g., Zhou et al., 2012; Diamond, 2013; Morrison & Grammer, 2016; Nigg, 2017), Duckworth and Kern (2011) conducted a meta-analysis examining the convergent validity of measures of self-regulation based on 282 independent samples. Specifically, they examined self-report surveys, informant-report surveys, cognitive tasks assessing executive function, and cognitive tasks assessing delay discounting. Duckworth and Kern (2011) concluded that the measures demonstrated moderate convergent validity overall, though the correlations among the measures showed substantial variability. Furthermore, they found much stronger evidence of convergent validity for the surveys than for the cognitive tasks.

To facilitate the integration of cross-disciplinary measures of self-regulation, we provide a database with the product-moment and distance correlations among 60 cross-disciplinary measures of self-regulation (23 self-report surveys, 37 cognitive tasks; see Eisenberg et al., 2018). As part of a larger project using these data, Eisenberg et al. (2019) investigated the dimensionality of these surveys and cognitive tasks via exploratory factor analysis, and Enkavi et al. (2019) evaluated the test-retest reliability of these surveys and cognitive tasks. Providing this database of product-moment and distance correlations is intended to complement previous work examining the convergent and divergent validity of these measures, including the meta-analysis conducted by Duckworth and Kern (2011). Whereas the correlations presented by Duckworth and Kern (2011) are based on summary statistics pooled across multiple publications and samples of participants (with each sample completing a subset of measures), the correlations presented in this paper are based on a single sample of participants who completed an extensive battery of surveys and cognitive tasks. We provide both product-moment correlations, which measure the strength and direction of the linear association between two variables, and distance correlations, which measure *any* dependence (i.e., linear and nonlinear associations) between two variables. These correlations can help assess convergent and divergent validity across the surveys and cognitive tasks. The database also includes health and substance use measures, which can help establish criterion validity. Other potential uses of the database include the specification of Bayesian prior distributions, meta-analysis, and integrative data analysis, which we describe later in the paper.

Methods

Participants and procedures

Data were collected on Amazon's Mechanical Turk (MTurk), an online marketplace where participants are paid to complete HITs (human intelligence tasks). In recent years, MTurk has gained popularity in psychological research for providing access to a large pool of participants who are often more diverse in geographic location, demographics, and abilities compared to participants recruited from universities and the surrounding communities. Crump, McDonnell, and Gureckis (2013) found that the quality of data collected on MTurk was comparable to that of data collected in laboratories. Hauser and Schwarz (2016) found that MTurk participants passed more online attention checks than did undergraduate students, though Barends and de Vries (2019) noted that some MTurk participants actively search for attention checks while otherwise providing careless responses.

Adults between 18 and 50 years old who were living in the United States were invited to participate, though four participants reported being between 50 and 60 years old. Participants were given one week to complete an approximately 10-hour battery of surveys and cognitive tasks, which were deployed in a random order through Experiment Factory (Sochat et al., 2016). Participants who completed the battery were paid $60 plus an average of $10 in bonuses based on performance (range = $65 to $75). Payments were prorated for those who did not complete the battery. The study protocol and data collection were preregistered on the Open Science Framework (https://osf.io/amxpv/).

Participants who did not complete the battery (102 of 662 participants who enrolled) or failed quality checks for the cognitive tasks (38 of 560 participants, described below) were excluded, resulting in a sample of 522 participants. Of the 522 participants, 50.2% were female and 78.8% were non-Latino White. Participants were fairly diverse in age ($M = 33.63$, $SD = 7.88$, range = 20–59 years old), education (0.6% less than high school, 14.8% high school diploma, 40.8% some college, 36.8% bachelor's degree, 7.1% master's degree or higher), and household income (M = $48,084, SD = $31,253, maximum = $240,000).

Measures

Based on an extensive review of the literature from various disciplines, we identified, administered, and scored 23 surveys and 37 cognitive tasks that all putatively measure constructs within the nomological network of self-regulation.

Surveys

The 23 surveys administered are listed in Table 1 and can be viewed on Experiment Factory (https://expfactory.github.

Table 1. Self-report surveys administered.

Barratt Impulsivity Scale (BIS-11)
Behavioral Inhibition System/Behavioral Activation System (BIS/BAS)
Brief Self-Control Scale
Carstensen Future Time Perspective
Dickman's Functional & Dysfunctional Impulsivity
Domain-Specific Risk-Taking Scale (DOSPERT), Expected Benefits
Domain-Specific Risk-Taking Scale (DOSPERT), Risk Perception
Domain-Specific Risk-Taking Scale (DOSPERT), Risk-Taking
Emotion Regulation Questionnaire
Eysenck I-7 Impulsiveness and Venturesomeness
Five-Facet Mindfulness Questionnaire
Short Grit Scale
Mindful Attention and Awareness Scale
Multidimensional Personality Questionnaire, Control
Selection Optimization Compensation Scale
Short Self-Regulation Questionnaire
Stanford Leisure Time Activities Survey
Ten-Item Personality Questionnaire
Theories of Willpower Scale
Three-Factor Eating Questionnaire
UPPS-P Impulsive Behavior Scale
Zimbardo Time Perspective Inventory
Zuckerman's Sensation Seeking Scale

Note. Key references for these surveys are available at https://scienceofbehaviorchange.org/measures/.

io/experiments/). A description of each survey is available at https://scienceofbehaviorchange.org/measures/.

Cognitive tasks

The 37 cognitive tasks administered are listed in Table 2 and can be viewed on Experiment Factory (https://expfactory.github.io/experiments/). A description of each cognitive task is available at https://scienceofbehaviorchange.org/measures/ (also see Enkavi et al., 2019). Quality checks were applied to all cognitive tasks to ensure that (1) average response times were not unreasonably fast, (2) not too many responses were omitted, (3) accuracy was reasonably high, and (4) responses were sufficiently distributed (i.e., the participant did not press only a single key). The specific criteria differed across cognitive tasks, but in general we required that median response times were longer than 200 milliseconds, no more than 25% of responses were omitted, accuracy was higher than 60%, and no single response was given more than 95% of the time. Failing a quality check resulted in that participant's data being removed for the failed cognitive task. Participants who failed quality checks for four or more cognitive tasks (38 of 560 participants) were excluded from the sample entirely. For the stop signal tasks, probabilistic selection task, and two-step decision task, performance criteria served as additional quality checks. Failing a quality check based on performance criteria led to removal of that participant's data for the failed cognitive task.

Demographics and health

Participants reported their demographics, height and weight, substance use (alcohol, nicotine, marijuana, and other drugs), history of medical and neurological disorders, and mental health (Kessler Psychological Distress Scale, K6; Kessler et al., 2002). Eleven of these variables were selected for inclusion in the database: body mass index; six items from the K6 on feeling nervous, hopeless, restless or fidgety, depressed, everything was an effort, or worthless in the past 30 days; one item on tobacco use ("On average, how many cigarettes do you now smoke a day?"); and three items on alcohol use ("How many drinks containing alcohol do you have on a typical day when you are drinking?", "How often do you have six or more drinks on one occasion?", "How often during the last year have you found that you were not able to stop drinking once you had started?"). Because self-regulation has been implicated in a variety of risky health behaviors, these measures can be used to evaluate the criterion validity of the surveys and cognitive tasks (see Eisenberg et al., 2019).

Table 2. Cognitive tasks administered.

Adaptive N-Back	Local-Global Letter Task
Angling Risk Task	Motor Selective Stop Signal
Attention Network Task	Probabilistic Selection Task
Bickel Titrator	Psychological Refractory Period
Choice Reaction Time	Raven's Progressive Matrices
Cognitive Reflection Task	Recent Probes
Columbia Card Task, Cold	Shape Matching Task
Columbia Card Task, Hot	Shift Task
Dietary Decision Task	Simon Task
Digit Span	Simple Reaction Time
Directed Forgetting	Spatial Span
Discount Titrator	Stimulus Selective Stop Signal
Dot Pattern Expectancy	Stop Signal
Go-No Go	Stroop
Hierarchical Learning Task	Three-By-Two
Holt & Laury	Tower of London
Information Sampling Task	Two-Step Decision
Keep Track Task	Writing Task
Kirby	

Note. Key references for these cognitive tasks are available at https://scienceofbehaviorchange.org/measures/.

Calculation of product-moment and distance correlations

Product-moment and distance correlations were computed among the 66 variables derived from the 23 surveys, 138 variables derived from the 37 cognitive tasks, and 11 variables pertaining to health and substance use (total = 215 variables). These 215 variables resulted in ${}_{215}C_2 = 23{,}005$ (product-moment or distance) correlations computed using pairwise deletion. The product-moment correlations were estimated using the *cor* function in R, and the distance correlations were estimated using the *dcor* function available through the *energy* package in R (Rizzo & Székely, 2016). Whereas the product-moment correlation measures the strength and direction of the linear association between two variables X and Y, the distance correlation measures *any* dependence (i.e., linear and nonlinear associations) between X and Y (Székely, Rizzo, & Bakirov, 2007). The distance correlation ranges from 0 to 1 and equals 0 only if X and Y are independent. The appendix provides the formula for the distance correlation.

Results

The correlation database is available as supplemental material, and the data used to generate these correlations are available at https://github.com/IanEisenberg/Self_Regulation_Ontology/tree/master/Data. Figure 1 presents the

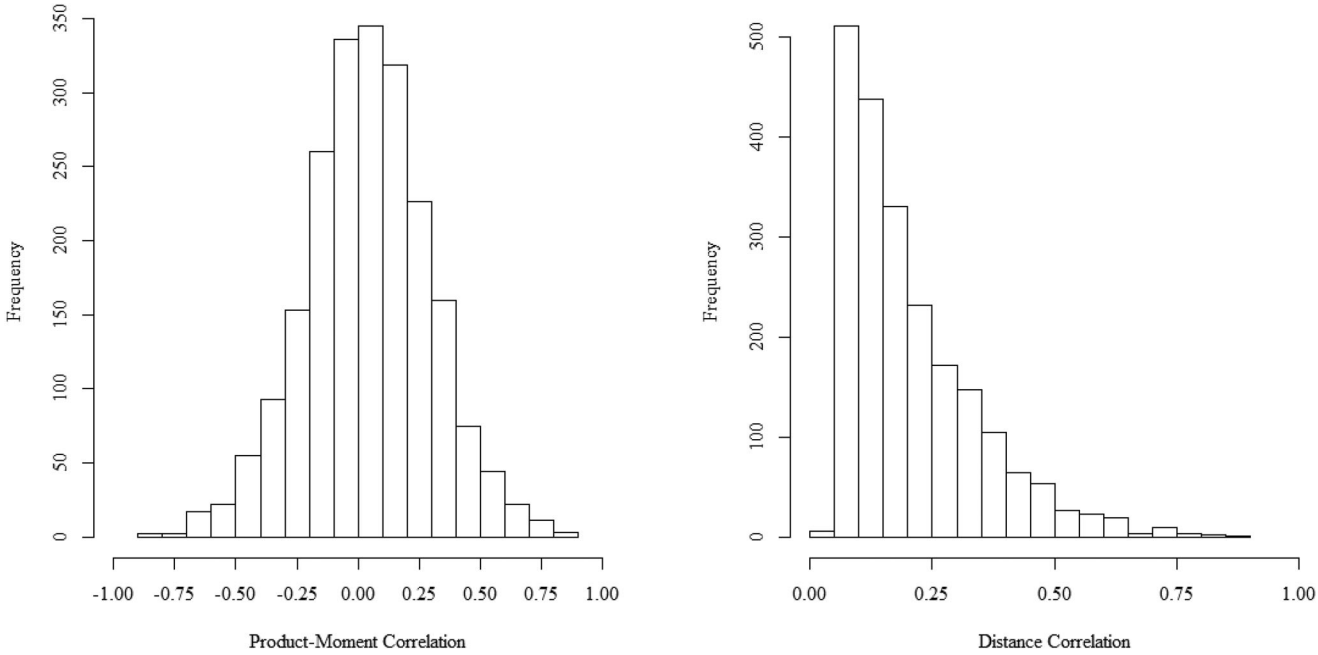

Figure 1. Distribution of product-moment and distance correlations among the 66 variables derived from the surveys.

distribution of product-moment and distance correlations among the 66 variables derived from the 23 surveys, and Figure 2 presents the distribution of product-moment and distance correlations among the 138 variables derived from the 37 cognitive tasks. As shown in Figures 1 and 2, the correlations showed substantial variability. Among the 66 survey variables, the absolute value of the product-moment correlations[1] ranged from .000 to .879 (M = .202), and the distance correlations ranged from .045 to .865 (M = .207). Among the 138 cognitive task variables, the absolute value of the product-moment correlations ranged from .000 to .974 (M = .087), and the distance correlations ranged from .028 to .956 (M = .121). The surveys demonstrated greater convergent validity than did the cognitive tasks based on both the product-moment and distance correlations (hereafter denoted r and $\mathscr{R}$, respectively).

A few of the most highly correlated variable pairs were from the Eysenck I-7, UPPS-P Impulsive Behavior Scale, and Zuckerman's Sensation Seeking Scale, likely because these surveys include identical or nearly identical items (e.g., "Would you enjoy water skiing?" on the Eysenck I-7; "I would enjoy water skiing" on the UPPS-P Impulsive Behavior Scale; and "I would like to take up the sport of water skiing, I would not like to take up water skiing" [forced choice] on Zuckerman's Sensation Seeking Scale). For example, the Eysenck I-7 venturesomeness subscale correlated with the UPPS-P sensation seeking subscale at r = .879 ($\mathscr{R}$ = .865) and Zuckerman's thrill seeking subscale at r = .865 ($\mathscr{R}$ = .840), and the UPPS-P sensation seeking subscale correlated with Zuckerman's thrill seeking subscale at r = .831 ($\mathscr{R}$ = .807). Other highly correlated variable pairs were derived from the same cognitive task and reflect different ways to score the same data from the cognitive task.

Overall, variables derived from the surveys were not highly correlated with variables derived from the cognitive tasks. For the $66 \times 138 = 9{,}108$ pairs involving one survey variable and one cognitive task variable, the absolute value of the product-moment correlations ranged from .000 to .271 (M = .049), and the distance correlations ranged from .028 to .241 (M = .085). The two strongest correlations between one survey variable and one cognitive task variable were as follows. Expected benefits from engaging in unethical behaviors (from the Domain-Specific Risk-Taking Scale [DOSPERT]; e.g., taking questionable deductions on your income tax return) and motivation on fixed win trials of the Information Sampling Task were correlated at r = .271 ($\mathscr{R}$ = .241). The Behavioral Inhibition System/Behavioral Activation System (BIS/BAS) drive subscale and Dot Pattern Expectancy d' were correlated at r = −.246 ($\mathscr{R}$ = .241). Many of the remaining correlations between one survey variable and one cognitive task variable were much lower than expected. For example, the Eysenck I-7 impulsiveness scale (e.g., "Do you generally do and say things without stopping to think?") only weakly correlated with the Tower of London planning time (r = −.023, $\mathscr{R}$ = .070), average move time (r = −.016, $\mathscr{R}$ = .065), number of extra moves (r = .063, $\mathscr{R}$ = .082), and number of optimal solutions (r = −.064, $\mathscr{R}$ = .092).

To illustrate the database's utility for assessing criterion validity, we highlight variables correlated with body mass index and alcohol use. Body mass index most strongly correlated with the emotional eating (r = .377, $\mathscr{R}$ = .371) and uncontrolled eating (r = .291, $\mathscr{R}$ = .297) subscales of the Three-Factor Eating Questionnaire, followed by the Brief Self-Control Scale (r = −.266, $\mathscr{R}$ = .281) and Stanford

[1]The absolute value was used to report the range and mean of the product-moment correlations because not all of the survey and cognitive task variables were scored in the same direction. That is, higher scores could represent greater self-regulation for some survey and cognitive task variables but lower self-regulation for others.

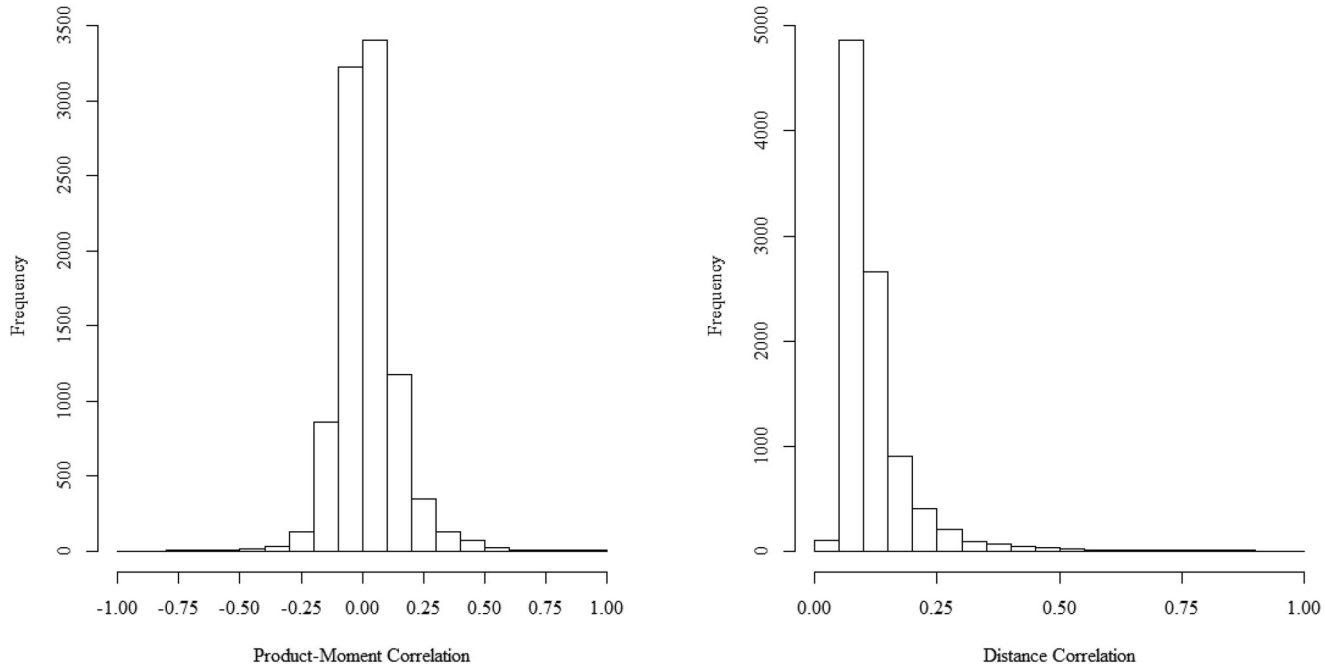

Figure 2. Distribution of product-moment and distance correlations among the 140 variables derived from the cognitive tasks.

Leisure Time Activities Survey (one item on physical activity; $r = -.209$, $\mathscr{R} = .241$). However, body mass index only weakly correlated with health sensitivity ($r = -.135$, $\mathscr{R} = .136$) and taste sensitivity ($r = -.042$, $\mathscr{R} = .074$) from the Dietary Decision Task. All three alcohol use items most strongly correlated with the disinhibition subscale of Zuckerman's Sensation Seeking Scale ($r = .435$, $\mathscr{R} = .436$ for number of drinks when drinking; $r = .514$, $\mathscr{R} = .492$ for having six or more drinks on one occasion; $r = .310$, $\mathscr{R} = .295$ for unable to stop drinking). This is likely due in part to forced choice items such as "Heavy drinking usually ruins a party because some people get loud and boisterous, Keeping the drinks full is the key to a good party" and "I feel best after taking a couple of drinks, Something is wrong with people who need liquor to feel good."

Overall, the distance correlations were consistent with the product-moment correlations. However, the product-moment and distance correlations differed by |.100| or more for 177 (0.8%) of the 23,005 variable pairs. Figure 3 provides scatterplots for two variable pairs—one where the distance correlation was .212 points *higher* in magnitude than the corresponding product-moment correlation and one where the distance correlation was .130 points *lower* in magnitude than the corresponding product-moment correlation. For the scatterplot on the left, Choice Reaction Time non-decision time showed greater variability at lower levels of Stimulus Selective Stop Signal non-decision time (both estimated from a drift diffusion model). This fanning out of scores may explain why the distance correlation (.235) was higher in magnitude than the corresponding product-moment correlation (−.022). For the scatterplot on the right, a few outliers appear to be driving up the product-moment correlation, such that the distance correlation (.257) was lower in magnitude than the corresponding product-moment correlation (−.388). When excluding the two cases with speeding scores less than −300, the product-moment and distance correlations were much less discrepant ($r = -.243$, $\mathscr{R} = .217$).

Discussion

Researchers across various disciplines have recognized the importance of self-regulation, but limited cross-talk among these researchers has stymied efforts to better understand and measure self-regulation. To facilitate the integration of cross-disciplinary measures of self-regulation, we provide a database with the product-moment and distance correlations among 60 cross-disciplinary measures of self-regulation (23 self-report surveys, 37 cognitive tasks) and measures of health and substance use based on 522 participants. The correlations showed substantial variability, though the surveys demonstrated greater convergent validity than did the cognitive tasks. Although we used overlapping (and not identical) sets of measures, our results coincide with those of Duckworth and Kern (2011). That is, Duckworth and Kern's (2011) meta-analysis similarly demonstrated greater convergent validity among the surveys (average $r = .50$ and .54 for self-report and informant-report, respectively) than among the cognitive tasks (average $r = .15$ and .21 for executive function and delay discounting, respectively).

Given that we selected cognitive tasks that all putatively measure constructs within the nomological network of self-regulation, poor convergent validity among the cognitive tasks may be due to poor construct validity of the cognitive tasks. That is, some of the cognitive tasks may not measure what they purport to measure. Alternatively, poor convergent validity among the cognitive tasks may be due to greater measurement error or difficulties with the purity of cognitive tasks. Cognitive tasks make demands on several cognitive processes not intended to be measured, thus

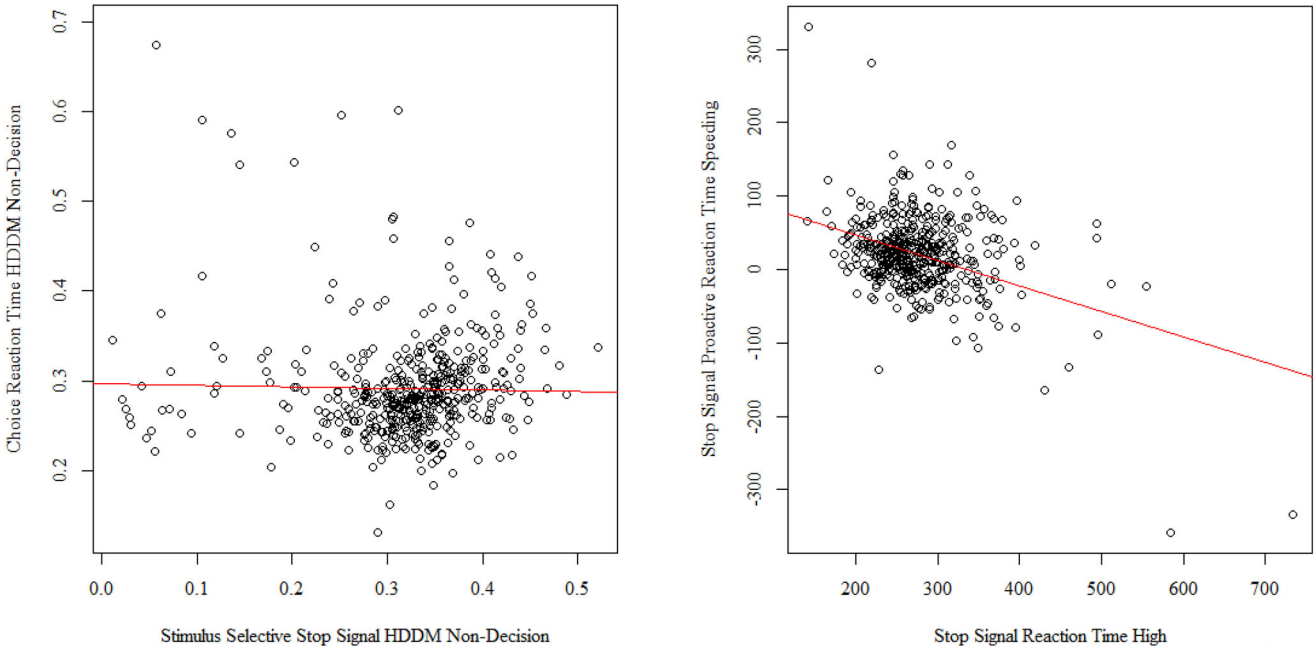

Figure 3. Scatterplots of variable pairs with discrepant product-moment and distance correlations. The ordinary least squares regression line is provided for each scatterplot. For the variable pair on the left, the distance correlation (.235) was .212 points *higher* in magnitude than the corresponding product-moment correlation (−.022). For the variable pair on the right, the distance correlation (.257) was .130 points *lower* in magnitude than the corresponding product-moment correlation (−.388).

calling into question the ability of cognitive tasks to isolate and validly measure a single cognitive process (Rabbitt, 1997). As an exploratory analysis, we estimated partial correlations among the cognitive task variables while controlling for participants' fluid intelligence as measured by Raven's Progressive Matrices (Raven, 1998). Conceivably, fluid intelligence may affect participants' performance on many of the cognitive tasks included in this study, though controlling for participants' fluid intelligence did not meaningfully change the relations among the cognitive task variables (M = .073 for the absolute value of the partial correlations).

Variables derived from the cognitive tasks were also not highly correlated with variables derived from the surveys. Astonishingly, the *maximum* product-moment correlation between one survey variable and one cognitive task variable was .271 despite examining 9,108 variable pairs. Controlling for participants' fluid intelligence did not meaningfully change the relations among the survey and cognitive task variables (M = .047 and maximum = .275 for the absolute value of the partial correlations). Similarly, in Duckworth and Kern's (2011) meta-analysis, variables derived from the surveys only weakly correlated with variables derived from the cognitive tasks (average r = .10 and .15 between self-report surveys and cognitive tasks assessing executive function and delay discounting, respectively). These weak correlations challenge the notion that the surveys and cognitive tasks measure the same construct. One possibility is that the surveys measure *trait* self-regulation whereas the cognitive tasks do not. Relative to the cognitive tasks, the surveys can more easily assess trait self-regulation by asking participants to consider their usual behaviors or abilities over time (e.g., "I usually think before I act" from the Short Self-Regulation Questionnaire, "I am able to work effectively toward long-term goals" from the Brief Self-Control Scale). Consistent with this hypothesis, in this sample, Enkavi et al. (2019) found much higher test-retest reliability for the surveys than for the cognitive tasks when repeated an average of 111 days later (median intraclass correlation = .674 and .311, respectively). More broadly, our results coincide with previous research suggesting that surveys and cognitive tasks often provide unique information when assessing psychological constructs (e.g., Meyer et al., 2001). Reconciling information gleaned from these two methods remains a challenge.

Finally, the health and substance use measures included in the database correlated more strongly with variables derived from the surveys than with those derived from the cognitive tasks. Although this pattern of results provides stronger evidence for the criterion validity of the surveys, other criteria not included in the database may be more strongly predicted by performance on the cognitive tasks. Furthermore, the surveys and health and substance use measures were all self-reported, which may inflate their correlations.

The database has several potential uses beyond our present use of examining the convergent validity of a wide array of surveys and cognitive tasks. First, the database can help researchers select a small subset of surveys and/or cognitive tasks to optimally assess self-regulation. Duckworth and Kern (2011) recommended administering both surveys and cognitive tasks and aggregating across multiple measures to reduce measurement error. Similarly, other researchers found that latent factors or summary measures combining across multiple cognitive tasks showed substantially greater test-retest reliability than the individual cognitive tasks (e.g., Miyake et al., 2000; Eisenberg et al., 2019). Although two highly correlated measures do not necessarily

assess the same construct, administering a second measure that is perfectly (or almost perfectly) correlated with the first provides little to no additional information. In our sample, variable pairs with very strong correlations (i.e., $|r| \geq .800$) were mostly limited to those derived from the same cognitive task or from surveys with identical or nearly identical items, though the UPPS-P lack of perseverance subscale correlated at $r = -.809$ ($\mathscr{R} = .772$) with the Short Grit Scale and at $r = -.802$ ($\mathscr{R} = .767$) with the Short Self-Regulation Questionnaire. Given these strong correlations, researchers generally should not pair these surveys to reduce the battery length and response burden. Furthermore, although surveys are easier and faster to administer than cognitive tasks, cognitive tasks provide added utility. For example, whereas surveys may suffer from response bias, performance on cognitive tasks is "difficult if not impossible to fake" (Duckworth & Kern, 2011, p. 266). When defining latent factors or summary measures based on multiple cognitive tasks, we would caution researchers to carefully consider the theoretical basis for these latent factors or summary measures and to critically review the fit of their measurement models given the weak correlations observed in this sample. These same considerations would apply when attempting to define latent factors or summary measures based on a combination of surveys and cognitive tasks. As noted previously, the surveys and cognitive tasks may not measure the same construct, such that combining them to form a single latent factor or summary measure would be inappropriate.

Second, the correlations may be used to specify Bayesian prior distributions. Although a review of the Bayesian framework is beyond the scope of this paper, briefly, parameters are treated as random in the Bayesian framework, not fixed as in the frequentist framework. A posterior distribution is defined for the parameter of interest by combining a user-specified prior distribution for the parameter with a model for the observed data. The prior distribution quantifies the user's prior beliefs about the parameter based on his or her expertise, results from a pilot investigation, other publications, and/or meta-analyses. A diffuse prior distribution assigns relatively little weight to the researcher's prior beliefs about the parameter, though more informative prior distributions can be specified depending on the availability of relevant findings. For example, Muthén and Asparouhov (2012) recently discussed the utility of a Bayesian approach to factor analysis. When adopting a frequentist approach, we can perform either exploratory factor analysis, which freely estimates all possible loadings, or confirmatory factor analysis, which constrains all cross-loadings to zero based on a theorized factor structure for the data. A Bayesian approach to factor analysis provides much more flexibility than a frequentist approach because we can encode prior beliefs that are stronger than those from exploratory factor analysis but weaker than those from confirmatory factor analysis (Levy & Mislevy, 2016). Specifically, Muthén and Asparouhov (2012) proposed assigning highly informative prior distributions to cross-loadings such that their posterior distributions are pulled toward zero but are not constrained to zero. The database provided in this paper can help researchers specify informative prior distributions for cross-loadings and other parameters when investigating the factor structure of constructs within the nomological network of self-regulation.

Finally, the database can be used when conducting a meta-analysis (either in the frequentist or Bayesian framework) or integrative data analysis. Integrative data analysis refers to pooling raw data across multiple samples of participants and fitting models to the pooled data (Curran & Hussong, 2009). Synthesizing different measures administered to different samples (referred to as measure harmonization) has been recognized as an important challenge and possible threat to the internal validity of integrative data analysis (Brincks et al., 2018). Because the database includes 60 cross-disciplinary surveys and cognitive tasks measuring self-regulation, it can help link samples that each complete a subset of these surveys or cognitive tasks.

By examining the relations among 60 cross-disciplinary measures of self-regulation, we aimed to provide a resource for addressing widespread confusion regarding the constructs within the nomological network of self-regulation and the measures used to assess these constructs. Given the importance of self-regulation, further efforts are needed to promote cross-disciplinary integration and cross-talk among researchers interested in self-regulation.

Funding

This work was supported by the National Institutes of Health Science of Behavior Change Common Fund under Grant UH2DA041713 administered by the National Institute on Drug Abuse.

ORCID

Gina L. Mazza http://orcid.org/0000-0002-5305-6193
Heather L. Smyth http://orcid.org/0000-0001-7865-4096
Sunny Jung Kim http://orcid.org/0000-0002-6405-162X
Felix Muniz http://orcid.org/0000-0002-2501-6469
Emily A. Scherer http://orcid.org/0000-0002-6166-4455
David P. MacKinnon http://orcid.org/0000-0003-0866-6010

References

Barends, A. J., & de Vries, R. E. (2019). Noncompliant responding: Comparing exclusion criteria in MTurk personality research to improve data quality. *Personality and Individual Differences*, *143*, 84–89. doi:10.1016/j.paid.2019.02.015

Bickel, W. K., Odum, A. L., & Madden, G. J. (1999). Impulsivity and cigarette smoking: Delay discounting in current, never, and ex-smokers. *Psychopharmacology*, *146*(4), 447–454. doi:10.1007/PL00005490

Bogg, T., & Roberts, B. W. (2004). Conscientiousness and health-related behaviors: A meta-analysis of the leading behavioral contributors to mortality. *Psychological Bulletin*, *130*(6), 887–919. doi:10.1037/0033-2909.130.6.887

Brincks, A., Montag, S., Howe, G. W., Huang, S., Siddique, J., Ahn, S., … Brown, C. H. (2018). Addressing methodologic challenges and minimizing threats to validity in synthesizing findings from individual-level data across longitudinal randomized trials. *Prevention Science*, *19*(S1), 60–73. doi:10.1007/s11121-017-0769-1

Crump, M. J. C., McDonnell, J. V., & Gureckis, T. M. (2013). Evaluating Amazon's Mechaninal Turk as a tool for experimental

behavioral research. *PLOS One.*, *8*(3), e57410–18. doi:10.1371/journal.pone.0057410

Curran, P. J., & Hussong, A. M. (2009). Integrative data analysis: The simultaneous analysis of multiple data sets. *Psychological Methods*, *14*(2), 81–100. doi:10.1037/a0015914

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64*(1), 135–168. doi: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*(3), 259–268. doi:10.1016/j.jrp.2011.02.004

Eisenberg, I. W., Bissett, P. G., Canning, J. R., Dallery, J., Enkavi, A. Z., Gabrieli, S. W., … Poldrack, R. A. (2018). Applying novel technologies and methods to inform the ontology of self-regulation. *Behaviour Research and Therapy*, *101*, 46–57. doi:10.1016/j.brat.2017.09.014

Eisenberg, I. W., Bissett, P. G., Enkavi, A. Z., Li, J., MacKinnon, D. P., Marsch, L. A., & Poldrack, R. A. (2019). Uncovering mental structure through data-driven ontology discovery. *Nature Communications*, *10*(1), 1–13. doi:10.1038/s41467-019-10301-1

Enkavi, A. Z., Eisenberg, I. W., Bissett, P. G., Mazza, G. L., MacKinnon, D. P., Marsch, L. A., & Poldrack, R. A. (2019). A large-scale analysis of test-retest reliabilities of self-regulation measures. *Proceedings of the National Academy of Sciences*, *116*(12), 5472–5477. doi:10.1073/pnas.1818430116

Eisenberg, N. (2017). Commentary: What's in a word (or words)—on the relations among self-regulation, self-control, executive functioning, effortful control, cognitive control, impulsivity, risk-taking, and inhibition for developmental psychopathology—reflections on Nigg. *Journal of Child Psychology and Psychiatry*, *58*(4), 384–386. (2017). doi:10.1111/jcpp.12707

Hauser, D. J., & Schwarz, N. (2016). Attentive Turkers: MTurk participants perform better on online attention checks than do subject pool participants. *Behavior Research Methods*, *48*(1), 400–407. doi:10.3758/s13428-015-0578-z

Kern, M. L., & Friedman, H. S. (2008). Do conscientious individuals live longer? A quantitative review. *Health Psychology*, *27*(5), 505–512. doi:10.1037/0278-6133.27.5.505

Kessler, R. C., Andrews, G., Colpe, L. J., Hiripi, E., Mroczek, D. K., Normand, S.-L T., … Zaslavsky, A. M. (2002). Short screening scales to monitor population prevalences and trends in non-specific psychological distress. *Psychological Medicine*, *32*(6), 959–976. doi:10.1017/S0033291702006074

Levy, R., & Mislevy, R. J. (2016). *Bayesian psychometric modeling*. Boca Raton, FL: Chapman & Hall/CRC.

Meyer, G. J., Finn, S. E., Eyde, L. D., Kay, G. G., Moreland, K. L., Dies, R. R., … Reed, G. M. (2001). Psychological testing and psychological assessment: A review of evidence and issues. *American Psychologist*, *56*(2), 128–165. doi:10.1037/0003-066X.56.2.128

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*(1), 49–100. doi:10.1006/cogp.1999.0734

Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H., … Caspi, A. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proceedings of the National Academy of Sciences of Sciences*, *108*(7), 2693–2698. doi:10.1073/pnas.1010076108

Morrison, F. J., & Grammer, J. K. (2016). Conceptual clutter and measurement mayhem: Proposals for cross-disciplinary integration in conceptualizing and measuring executive function. In J. A. Griffin, P. McCardle, & L. Freund (Eds.), *Executive function in preschool-age children: Integrating measurement, neurodevelopment, and translational research* (pp. 327–348). Washington, DC: American Psychological Association. doi:10.1037/14797-015

Muthén, B., & Asparouhov, T. (2012). Bayesian structural equation modeling: A more flexible representation of substantive theory. *Psychological Methods*, *17*(3), 313–335. doi:10.1037/a0026802

National Institutes of Health (2015, January 8). *Assay development and validation for self-regulation targets*. Retrieved from https://grants.nih.gov/grants/guide/rfa-files/RFA-RM-14-020.html.

Nigg, J. T. (2017). Annual Research Review: On the relations among self-regulation, self-control, executive functioning, effortful control, cognitive control, impulsivity, risk-taking, and inhibition for developmental psychopathology. *Journal of Child Psychology and Psychiatry*, *58*(4), 361–383. doi:10.1111/jcpp.12675

Rabbitt, P. (1997). Introduction: Methodologies and models in the study of executive function. In P. Rabbitt (Ed.), *Methodology of frontal and executive function* (pp. 1–37). East Sussex: Psychology Press.

Raven, J. C. (1998). *Raven's progressive matrices and vocabulary scales*. Oxford: Oxford Pyschologists Press.

Rizzo, M. L., & Székely, G. J. (2016). *E-statistics: Multivariate inference via the energy of data*. R package version 1.7-6. https://CRAN.R-project.org/package=energy

Sochat, V. V., Eisenberg, I. W., Enkavi, A. Z., Li, J., Bissett, P. G., & Poldrack, R. A. (2016). The Experiment Factory: Standardizing behavioral experiments. *Frontiers in Psychology*, *7*, 1–9. doi:10.3389/fpsyg.2016.00610

Székely, G. J., Rizzo, M. L., & Bakirov, N. K. (2007). Measuring and testing dependence by correlation of distances. *The Annals of Statistics*, *35*(6), 2769–2794. doi:10.1214/009053607000000505

Welsh, M., & Peterson, E. (2014). Issues in the conceptualization and assessment of hot executive functions in childhood. *Journal of the International Neuropsychological Society*, *20*(2), 152–156. doi:10.1017/S1355617713001379

Zhou, Q., Chen, S. H., & Main, A. (2012). Commonalities and differences in the research on children's effortful control and executive function: A call for an integrated model of self-regulation. *Child Development Perspectives*, *6*(2), 112–121. doi:10.1111/j.1750-8606.2011.00176.x

Appendix

Distance correlation formula

The formulas and notation used throughout this appendix are based on Székely et al. (2007). The distance correlation of X and Y is defined as

$$\mathcal{R}(X,Y) = \sqrt{\frac{\mathcal{V}^2(X,Y)}{\mathcal{V}^2(X)\mathcal{V}^2(Y)}} \tag{A1}$$

where $\mathcal{V}^2(X,Y)$ is the squared distance covariance of X and Y, $\mathcal{V}^2(X)$ is the distance variance of X, $\mathcal{V}^2(Y)$ is the distance variance of Y, and $\mathcal{V}^2(X)\mathcal{V}^2(Y) > 0$. The distance covariance of X and Y is defined as

$$\mathcal{V}(X,Y) = \sqrt{\frac{\sum_{k,l=1}^{n} A_{kl}B_{kl}}{n^2}} \tag{A2}$$

where A and B are $n \times n$ matrices and n denotes the number of participants. Each element in A equals $a_{kl} - \overline{a}_{k.} - \overline{a}_{.l} + \overline{a}_{..}$ where $a_{kl} = |X_k - X_l|_p$, $\overline{a}_{k.} = \frac{\sum_{l=1}^{n} a_{kl}}{n}$ and $\overline{a}_{.l} = \frac{\sum_{k=1}^{n} a_{kl}}{n}$ are marginal means, $\overline{a}_{..} = \frac{\sum_{k,l=1}^{n} a_{kl}}{n^2}$ is the grand mean, and $k, l = 1, 2, \ldots, n$. Similarly, each element in B equals $b_{kl} - \overline{b}_{k.} - \overline{b}_{.l} + \overline{b}_{..}$ where $b_{kl} = |Y_k - Y_l|_q$, $\overline{b}_{k.} = \frac{\sum_{l=1}^{n} b_{kl}}{n}$ and $\overline{b}_{.l} = \frac{\sum_{k=1}^{n} b_{kl}}{n}$ are marginal means, $\overline{b}_{..} = \frac{\sum_{k,l=1}^{n} b_{kl}}{n^2}$ is the grand mean, and $k, l = 1, 2, \ldots, n$. The distance variance of X is defined as

$$\mathcal{V}^2(X) = \sqrt{\frac{\sum_{k,l=1}^{n} A_{kl}^2}{n^2}} \tag{A3}$$

and the distance variance of Y is defined as

$$\mathcal{V}^2(Y) = \sqrt{\frac{\sum_{k,l=1}^{n} B_{kl}^2}{n^2}} \tag{A4}$$

where A and B have the same definitions as above.